NIST
National Institute of Standards and Technology
Technology Administration, United States Department of Commerce

NIST Technical Note 1500-9
Materials Reliability Series

Charpy Verification Program: Reports Covering 1989–2002

C. N. McCowan
T. A. Siewert
D.P. Vigliotti

NIST Technical Note 1500-9
Materials Reliability Series

Charpy Verification Program: Reports Covering 1989-2002

C. N. McCowan
T. A. Siewert
D. P. Vigliotti
Materials Reliability Division, Materials Science and Engineering Laboratory
National Institute of Standards and Technology

September 2003

U.S. DEPARTMENT OF COMMERCE
Donald L. Evans, Secretary

TECHNOLOGY ADMINISTRATION
Phillip J. Bond, Under Secretary for Technology

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
Arden L. Bement, Jr., Director

National Institute of Standards and Technology Technical Note 1500-9
Natl. Inst. Stand. Technol., Tech. Note 1500-8, 375 pages (September 2003)
CODEN:NTNOEF

U.S. GOVERNMENT PRINTING OFFICE
WASHINGTON: 2003

For sale by the Superintendent of Documents, U.S. Government Printing Office
Internet: bookstore.gpo.gov Phone: (202) 512-1800 Fax: (202) 512-2250
Mail: Stop SSOP, Washington, DC 20402-001

Preface

This is a compendium of various studies on Charpy impact test procedure and specimen preparation variables in which NIST staff has participated in the past decade or so. Some of these studies had detail sufficient to make them worthy of peer review and publication in archival journals, others were just simple studies of one or two process parameters, and have never been published. However, all of them still contain information that is useful to those who produced reference grade verification specimens or have an interest in understanding how the various processing and procedure parameters affect absorbed energy. The collection serves as a single, archival source for these reports and the data that they contain.

This Technical Note includes 23 presentations, grouped into sections on:

1. Overview and History
2. Materials and Heat Treatment
3. Specimen, Machine, and Procedure Effects
4. Statistical Evaluations of Charpy Impact Data

Since the late 1980s, NIST has furnished standard reference materials for use in verifying the performance of Charpy impact machine according to the requirements of ASTM Standard E 23. Some of these reports describe improvements in processing procedures for the specimens (for reduced scatter), while other reports indicate how changes in the testing procedures affect the impact energy.

Acknowledgment: The editor wishes to express his appreciation to many other NIST staff members and students who contributed to these studies and helped to run the verification program. These include James Alcorn, Erin Alexander, Jenniffer Caragol, Diane Cyr, Samantha Dimmick, Lisa Dirling, Cathleen Farrell, Matthew Kuhn, Leslie Leininger, Brian Marsh, Nicole Neumeyer, Emma Nicoletti, Jason Pepin, Joanna Perez, Alisha Rodriguez, Dominique Shepherd, Jesse Sycuro and Meigen Thomas.

C.N. McCowan
May 1, 2003

Foreword

The Materials Reliability Series of NIST Technical Notes are reports covering significant research accomplishments of the Materials Reliability Division. The Division develops measurement technologies that enable the producers and users of materials to improve the quality and consistency of materials, for nondestructive evaluation to assure quality of finished materials and products, and for materials evaluation to assure reliable performance.

This report is the ninth in the series. It documents our publications and studies on Charpy impact machines for the past decade. Previous reports in this series are:

Technical Note 1500-1	Tensile Testing of Thin Films: Techniques and Results, by D.T. Read, 1997
Technical Note 1500-2	Procedures for the Electron-Beam Moire Technique, by E.S. Drexler, 1998
Technical Note 1500-3	High-Energy, Transmission X-ray Diffraction for Monitoring Turbine-Blade Solidification, by D.W. Fitting, W.P. Dubé, and T.A. Siewert , 1998
Technical Note 1500-4	Nondestructive Characterization of Reactor Pressure Vessel Steels: A Feasibility Study, by H.I. McHenry and G.A. Alers, 1998
Technical Note 1500-5	Electron-Beam Moiré Technique: Advances, Verification, Application, by E.L. Drexler, 1998
Technical Note 1500-6	Constitutive Behavior Modeling of Steels Under Hot-Rolling Conditions, by Y.W. Cheng, 1999
Technical Note 1500-7	Structure-Property Relationships in Steel Produced in Hot-Strip Mills, by P.T. Purtscher, Y.W. Cheng, and C.N. McCowan, 1999
Technical Note 1500-8	Recommended Practice: Installing, Maintaining, and Verifying Your Charpy Impact Machine, by D.P. Vigliotti, T.A. Siewert, and C.N. McCowan, 2002

Contents

Part 1: Overview and History 1

The NIST Charpy V-notch Verification Program: Overview and Operating Procedures 3
The History and Importance of Impact Testing 43
For 100 Years, Notched Bar Impact Testing Standards Have Yielded Widespread Benefits 55
Impact Test Methods: Procedures and Analysis 67

Part 2: Materials and Heat Treatment 83

Evaluation of ABS Plastic Impact Verification Specimens 85
Performance Verification of Impact Machines for Testing Plastics 99
Standard and Experimental Heat Treatments for Low-Energy 4340 Steel Verification Specimens 115
Homogenization and Subcritical Annealing: 4340 Charpy V-notch Verification Specimens 151
Heat Treatment of NIST T-200 CVN Specimens 159
Evaluation of 11.1 mm Bar for Service in the Charpy V-notch Program..... 183
Investigation of Cracks in CharpyV-notch Specimens 189

Part 3: Specimen, Machine, and Procedure Effects 203

Report of Measurement of Three Charpy Impact Machines at the National Institute of Standards and Technology 205
Maintaining the Accuracy of Charpy Impact Machines 215
Effect of Reduced Notch Radius on Charpy Impact Energy 225
The Effect of Charpy V-Notch Striker Radii on the Absorbed Energy 231
Low Cost Lower Bound Toughness Measurements 243
The Role of Strike Marks on the Reproducibility of Charpy Impact Test Results 261
Effect of Surface Finish of Charpy Anvils and Striking Bits on Absorbed Energy 279
The Influence of Shear Lip Symmetry on the Absorbed Energy of Charpy Impact Specimens ... 287

Part 4: Statistical Evaluations of Charpy Impact Data 295

International Comparison of Impact Verification Programs 313
Consensus Values and Reference Values Illustrated by the Charpy Machine Certification Program 331
Proposed Changes to the Charpy V-notch Machine Certification Requirements 349
Uncertainty in Reference Values for the Charpy V-notch Program 362

Part 1: Overview and History

The NIST Charpy V-notch Verification Program: Overview and Operating Procedures

C.N. McCowan, T.A. Siewert, D.P. Vigliotti
NIST, Materials Reliability Division

This report documents the procedures used by NIST in the Charpy V-notch reference material program. It was prepared to provide outside observers with accurate and detailed information on how the Charpy verification program is conducted, and also to serve as the basis for an internal record that will be updated to reflect when and why changes were made to the program.

Keywords: Charpy V-notch test, impact testing, mechanical testing, NIST, reference material, verification program

1. Introduction

1.1 Background

Charpy impact testing is often specified as an acceptance test for structural materials, and companies performing acceptance tests are typically required to verify the performance of their impact machine periodically. The procedure for verifying the performance of Charpy impact machines has a physical part and an engineering part. The physical part covers the direct verification of the impact machine, through a detailed evaluation of the machine dimensions, alignment, etc. The engineering part covers the indirect verification of the machine performance, which entails breaking sets of Charpy impact reference specimens. The indirect verification procedure was added about 40 years ago, because the use of direct verification procedures alone could not explain some unacceptable differences among the results of the machines tested. Often these differences could be traced to interactions between the machine components and the specimens, and only testing with verification specimens could resolve these effects. NIST supplies the impact reference specimens used to indirectly verify the performance of machines according to ASTM Standard E 23. The procedures used to conduct this program are the focus of this report.

Originally, the U.S. Army (Watertown Arsenal, AMMRC) produced and distributed the reference specimens for the verification of impact machines in the United States. NIST took over the program from the Army in 1989, and Army personnel helped to transfer the Charpy machines and their evaluation procedures to NIST. The three Charpy machines owned by the Army, and now by NIST, have been defined in ASTM Standard E 23 as the "master Charpy impact machines" for the United States for more than 12 years. Some of these machines have been in the program for 30 years. Each year, the verification test results for approximately 1000 industrial machines are evaluated. If the results of the industrial machines agree with the results of the master machines within either 1.4 J or 5 %, the machines are certified for acceptance testing according to the requirements of ASTM Standard E 23.

1.2 Relationship to National and International Standards

ASTM Standard E 23 requires the indirect verification of Charpy V-notch impact machines annually. Also, the ASME *Boiler and Pressure Vessel Code*, many U.S. military procurement specifications, and several ISO Standards require Charpy impact machine verification. The European and Japanese verification programs provide traceability to national laboratories and some international agreements are written to require traceability to a national laboratory. The NIST verification specimens meet ASTM and ISO requirements for indirect verification testing of Charpy V-notch impact machines and also provide traceability to a national laboratory.

Currently there are four laboratories in the world that certify and distribute reference materials for the verification of Charpy impact machines: (1) The Institute for Reference Materials and Measurements (IRMM, Belgium), (2) Laboratoire National d'Essais (LNE, France), (3) The National Institute of Standards and Technology (NIST, USA), and (4) The National Research Laboratory of Metrology (NRLM, Japan). These four laboratories supply specimens to verify the performance of more than 2000 impact machines annually.

1.3 Industrial Needs Met by the Program

The primary purpose of the program is to provide U.S. industry a source for high-quality impact verification specimens. Because this program also offers an evaluation service for the results of verification tests, several additional benefits are provided to industry. There are direct benefits to the customer in the evaluation and interpretation of the test results by NIST, such as a verification letter that auditors acknowledge and rarely question. Indirect benefits, due to the centralized evaluation of all verification test results include: (1) some assurance that trading partners and competitors are reporting comparable impact values, and (2) a centralized database that can support arguments for or against changes to national and international impact standards that affect U.S. industries.

2. Program Design Philosophy and Scope

The NIST Charpy verification program is designed to provide a complete service for our customers. The program can be divided into three basic parts, as follows: (1) The production of the verification materials, which includes purchasing the raw materials, contracting for their heat treatment and machining, batch certifications of verification specimens, and distribution of the specimens. (2) The evaluation of verification test results, which entails evaluating the data and specimens received from our customers, entering the data into our database, writing verification letters that include specific remarks concerning the performance of the machines being evaluated, and communicating with customers by fax, phone, and email regarding the outcome of tests. (3) The evaluation of data in our database, which serves as a final quality-control tool on our verification specimens, and as a means to evaluate and track the performance of industrial machines.

The accurate and consistent certification of the absorbed energy for our verification materials is the central part of the program, upon which all else is based. As is the case for many

measurements, these factors are difficult to address, particularly consistency over long-time intervals. Our procedures are based on the fact that the master machines at NIST are the designated reference machines for the United States (by ASTM E 23), and by our own definition the average value of the three machines is correct.

Our primary control for the program is tracking the performance of each master machine relative to the others. A change in the performance of one machine initiates an evaluation of that machine and the measurement system in general. Although this approach is a practical solution to a complex problem, and clearly has shortcomings, it has provided a robust and stable base for our certification procedures over the last 12 years. Additional controls to evaluate the quality of the verification specimens, such as the testing of control specimens and constant monitoring of the average energy values from customer verification test results are also used to monitor the quality of the specimens. In retrospect, we find that over the last 12 years, pooled data from the three master machines has proven to be a reliable and reasonable target for measuring the performance of industrial impact machines.

3. Description of Equipment and Personnel

The impact machines used by NIST were purchased from three different commercial suppliers, not custom built at NIST, and so represent the machines used by industries around the world. This is true for most of the equipment used in the program.

3.1 Impact Energy Measurement

We have five Charpy V-notch impact machines that are used for the measurement of absorbed impact energy in the program:*

Machine #1 Tokyo Koki Seizosho, "C" type pendulum, S/N 878303
359 Joule Capacity, Reference Machine

Machine #2 Tinius Olsen, Model 74, "U" type pendulum, S/N 130005
358 Joule Capacity, Reference Machine

Machine #3 Satec, Model SI-1C, "U" type pendulum, S/N 1262
325 Joule Capacity, Reference Machine

Machine #4 Satec, Model SI-1K3, "U" type pendulum, S/N 1662
407 Joule Capacity, Research Machine

Machine #5 Tinius Olsen, Model 84, "U" type pendulum, S/N 165153
407 Joule Capacity, Research Machine

*Trade names and names of manufacturers are included in several places in this report to accurately describe NIST activities. Such inclusion neither constitutes or implies endorsement by NIST or by the U.S. government.

Machines 1, 2, and 3 are the primary reference machines (the master machines). Machines 4 and 5 have higher capacity, newer designs that will eventually replace the older machines (as needed, once we are certain of their stability). All of the impact machines are equipped with optical encoders and digital readouts. Machine 4 has an instrumented striker.

The master machines are used only for assigning certified energy values to lots of verification specimens, and for occasional participation in measurement development programs and international round robins. The two backup machines are used for research on conventional and instrumented Charpy V-notch testing.

3.2 Hardness Measurements

Measurements are made on a commercial hardness testing machine. The tester is linked to a personal computer that is used to acquire and file data for the tests. These data are processed to evaluate the hardness level, and the uniformity in the hardness, of our verification specimens.

3.3 Dimensional Measurements

The notch depth, radius, angle, and centering are measured on a commercial optical comparator (50X) prior to impact testing. The squareness is measured with a gage described in ASTM E 23. The overall specimen dimensions are measured with digital calipers. A second, older optical comparator is used as a backup system for dimensional measurements. Data from the optical comparators and the calipers are output to a personal computer.

3.4 Software

Software was developed by NIST personnel to help manage specific tasks that are routinely performed when evaluating customer test results or certifying a production lot of verification specimens.

The database of customer data and information is organized as follows: (1) A main panel that contains fields for the serial number, manufacturer, capacity, and pendulum design of the machine, along with customer information such as the company name, address, and contact person, and also contains a comment field and the pass/fail status from the last verification test made on the machine. (2) A data panel that contains fields for the serial number of the machine, record number, test evaluation date, initials of the NIST operator who evaluated the data, data fields for the energy data (for four energy levels), the series number of the lot tested, automatically calculated fields for the customer's average energy, the NIST reference value, the difference between the customer and NIST energy values, and a pass/fail status field. (3) A panel containing information on test companies (so address information is available to address letters to third parties who conduct verification tests for the customer). (4) A reference value panel that contains all of the certified energy values for our verification lots.

A word processing program is used to help write customer letters. The program uses macros and Boolean logic to construct the letters according to operator input at various prompts. All of the pertinent data concerning the customers' test results and address information are accessible in the

program, and in addition, approximately 60 "standardized comments" are available for selection by the operator to help construct the letter.

A NIST data program is also used to collect and calculate output from the hardness tester and the dimensional measurement equipment.

3.5 Personnel

Three people are involved with the Charpy verification program in Boulder. One specializes in the operation and maintenance of the impact machines, and typically handles all of the day-to-day operations of the program (customer evaluations and service, and pilot lot certifications). The second specializes in the issues relating to the materials used to produce the verification materials, and in standards governing impact verification testing (ASTM and ISO). The third oversees the program. All the personnel involved in the program are capable of filling in for the others, which provides adequate backup for the program.

4. Procurement Requirements for Verification Materials

Two materials are currently used to make the specimens for the indirect verification of Charpy impact machines. An AISI type 4340 steel is used to make specimens at the low-and-high energy levels. A type T-200 maraging steel is used to make specimens at the super-high energy level.

4.1 Type 4340 Steel

4.1.1 Compositional and melting requirements

We require AISI 4340 steel bars, from a single heat to minimize compositional and micro-structural variation. Because steel plants produce steel in different heat sizes (inherent to their facilities), we have tried to add some flexibility in our contracts by bracketing the quantity of the steel to be purchased. We prefer to purchase about 5000 kg (5 ton) heats. The bids are evaluated primarily on cost, but we also consider delivery time. The composition for the heat of type 4340 steel that NIST is currently using is given in **Table 1**.

Table 1. Composition of 4340 steel (mass %).

C	Si	Mn	Ni	Cr	Mo	S	P
0.4	0.28	0.66	1.77	0.83	0.28	0.001	0.004

The steel is required to be produced using a double-vacuum-melting procedure (vacuum-induction-melt vacuum-arc-remelt) and meet the compositional requirements of AISI-SAE alloy 4340. The steel must also meet the stricter requirements of AMS 6414, which describes steel production by a vacuum-melting procedure. In addition, we desire the phosphorus, sulfur, vanadium, niobium, titanium, and copper contents of the steel to be as low as possible. The

maximum concentrations (in mass %) allowed for these elements are P = 0.010, S = 0.005, V = 0.030, Nb = 0.005, Ti = 0.003, and Cu = 0.35.[1]

The composition is certified using standard analytical procedures (such as optical emission spectroscopy or x-ray fluorescence), and the equipment is calibrated by standards traceable to NIST. The composition is measured at the top and bottom remelted ingot. These two measurements must be included in the documentation with the order, meet our compositional requirements, and meet the limits on residuals given above. Deviations between the two measurements (in mass percent between the top and bottom of VAR ingot) can not exceed 0.020 for C, Si, and Mo; 0.090 for Mn; 0.030 for Cr; 0.040 for Ni; 0.002 for P; and 0.001 for S.

The compositions of three heats of 4340 steel we have used are given in **Table 2**. The current alloy in use, heat number E4261, was processed from six ingots and the label on the bar indicates the location (top or bottom) in the ingot from which it was produced. This heat yielded 8699 kg of bar stock from a gross ingot weight of 9221 kg.

4.1.2 **Product form**

The ingots are forged, hot rolled, then cold finished to 12.7 mm square bars (+3.8 mm, −0.0 mm) and annealed. The corner radius of the finished bars cannot exceed 0.76 mm. The maximum acceptable grain size is ASTM number 8. In other attributes (decarburization, surface condition, etc.), the steel must be suitable for use as 10 mm square Charpy V-notch specimens.

The bar is normalized at 950 °C, and hardened to approximately 35 Rockwell C (HRC). We will accept alternate heat-treating schedules by mutual agreement. Our goal here is to produce bars with a minimum of large carbides in the structure, the most uniform carbide precipitation possible, and a uniform hardness. The bar is required to be machine straightened (for twist and bow), and shipped in lengths of no less than 2 m and no more than 4 m.

Table 2. Composition of 4340 alloy, mass %.

Heat #	2397 B	E4261	E4871
Year	1990	1993	1998
C	0.42	0.40	0.43
Mn	0.75	0.67	0.70
P	0.008	0.004	0.004
S	0.001	0.001	0.002
Si	0.30	0.28	0.30
Cr	0.84	0.83	0.82
Ni	1.83	1.77	1.78
Mo	0.27	0.28	0.24
Cu	0.03	NA	0.09

4.1.3 **Packaging**

The bar is packaged in bundles identified with reference to the ingot position from which it was processed. This identification is used to limit the material used for a given pilot lot to a single ingot location, which reduces microstructural inhomogeneities between bars. The bundles must weigh less than 900 kg (4000 lb), which is the capacity of our fork lift truck.

[1] There has been some question whether this very low sulfur level is the optimum level. Internal data from one steel producer indicate that sulfur levels of 0.01 to 0.03 may help reduce the variation in impact toughness for 4340 steels.

4.2.1 **Compositional and melting requirements**

We require double-vacuum-melted 18 Ni maraging steel bars. The steel must be of a single heat and the ingot(s) must be adequately forged prior to rolling to minimize compositional and microstructural variation in the final products. The steel must be produced using a vacuum-induction-melt vacuum-arc-remelt (VIM/VAR) procedure, and meet the nominal compositional requirements given in **Table 3**:

Table 3. Type T-200 steel (mass percent).

Ni	Mo	Ti	Al	Si, max	Mn, max	C, max	S, max	Co, max	P, max
18.5	3.0	0.7	0.1	0.1	0.1	0.01	0.01	0.5	0.01

The composition must be certified using standard analytical procedures, using equipment calibrated by standards traceable to NIST. The composition is measured at the top and bottom of the ingot. These two measurements are included in the documentation with the order, and must meet the requirements given above within reasonable tolerances for an 18 Ni maraging steel. If the presence of any residual elements (not included in the requirements above) are expected for the alloy, a maximum allowable concentration for this element must be agreed upon. Deviations between the two measurements (in mass percent between the top and bottom of VAR ingot) can not exceed those expected for high-quality VIM/VAR 18 Ni ingots (by current steel making standards).

Information on the current T-200 material we are using is given in **Table 4**. The three columns of data represent results of samples taken from the top and bottom of the VAR ingot, and the melt used to make the ingots. The alloy was melted in a vacuum induction furnace and cast into an electrode mold approximately 432 mm in diameter which weighed 3630 kg (4 ton). The electrode was remelted into an ingot with a diameter of 508 mm, which was cropped (3175 kg) and 100 % conditioned prior to chemical analysis (top and bottom). The 508 mm ingot was forged to a 432 mm octagon, then to 350 mm square, then to 250 mm square, and cut into 6 equal lengths. The 250 mm square was then forged to 152 mm square (billet) and air cooled. Each 152 mm billet was cut into three lengths, resulting in a total of 18. These 18 billets are coded and each bundle of bar that NIST received is from a single billet. The 152 mm billets were direct rolled on a mill to 57 mm and cut to lengths of approximately 660 mm prior to final rolling.

Table 4. T-200 composition.

	Top of Ingot	Bottom of Ingot	Melt
C	0.003	0.002	0.004
Mn	0.03	0.03	0.02
P	0.007	0.007	0.005
S	0.003	0.003	0.003
Si	0.01	0.01	0.01
Cr	0.20	0.21	0.20
Ni	18.79	18.77	18.77
Mo	3.01	2.97	2.89
W	0.02	0.01	<0.01
V	0.01	0.01	0.01
Co	0.47	0.47	0.51
Cu	0.01	0.01	0.01
Ti	0.79	0.78	0.80
Al	0.11	0.11	0.128
B	0.0008	0.0008	<0.002
Zr	0.005	0.005	NA

4.2.2 **Product form**
The heat is processed to 12.7 mm (+3.8 mm, −0.0 mm) square bar. The corner radius of the finished bars cannot exceed 0.76 mm. The maximum average grain size accepted is ASTM number 10. In other attributes (surface condition, etc), the steel is required to be suitable for use as 10 mm square Charpy V-notch specimens. The bars are delivered in the as-rolled condition. The bar is machine straightened (for twist and bow), and shipped in lengths of no less than 2 m and no more than 4 m.

4.2.3 **Packaging**
The bar is required to be packaged in bundles identified with reference to the ingot and which portion of the ingot it was processed from. If possible, the bundles consist of bar rolled from individual billets used for the rolling operation. The bundles must weigh less than 1815 kg (2 ton).

5. Specimen Production

5.1 Heat Treatment

5.1.1 Type 4340 steel
The 4340 steel is heat treated to produce low- and high-energy verification specimens. Typically, as indicated in **Figure 1**, low energy levels are attained by tempering at temperatures between 300 and 400 °C. The high energy specimens are tempered near 600 °C. The microstructure of the specimens must be 100 % tempered martensite.

The heat treatments originally recommended by the Army Materials Technology Laboratory are shown in **Table 5**.

Although the heat treatment of 4340 steel is straightforward for most commercial applications, it is not easy to produce the quality required for the impact verification specimens, particularly for production lots of approximately 1200 specimens.

One reason for this is that the transition behavior, shown in **Figure 2**, is not ideal for 4340 steel: at −40 °C the upper shelf of the high-energy specimens, and the lower shelf of the low-energy specimens are not flat. This can result in increasing the scatter during testing. Added to this are the effects of slight differences in heat treating between specimens, slight inhomogeneities in the steel, and other considerations. So, for our case, where a maximum range in hardness of less than 0.5 HRC is needed for a production lot, slight differences in the thermal history of the specimens can quickly present problems. Our experience has shown that the heat treatments recommended by the Army can give good results for small lot sizes. For example, we had two heat treating shops follow these recommendations to produce two low-energy lots for impact testing. No additional heat treatment specifications were added to our instructions, so different quench oils, etc. were used by the shops.

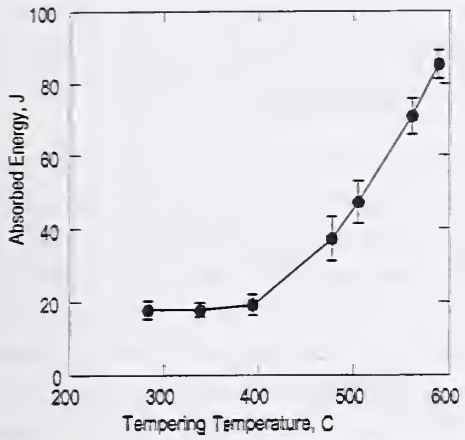


Figure 1. Data for 4340 verification material, 2001.

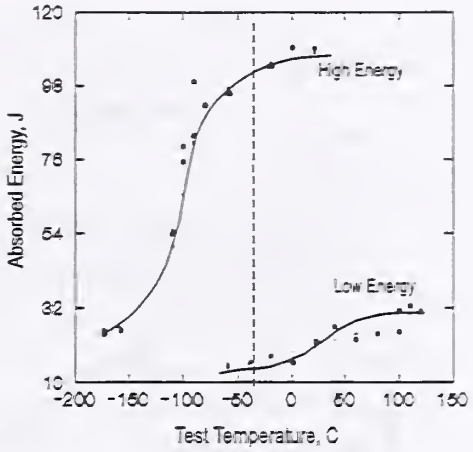


Figure 2. Transition curves for 4340 steel that has been heat treated for low- and high-energy verification specimens.

Table 5. Example heat treatments for low- and high-energy level type 4340 impact specimens.

Low-energy specimens, hardness 46 HRC ±1 HRC	High-energy specimens, hardness 32 HRC ±1 HRC
Normalize 900 °C (1650 °F) for 1 h, air cool	Normalize 900 °C (1650 °F) for 1 h, air cool
Harden 871 °C (1600 ° F) for 1 h, oil quench	Harden 871 °C (1600 ° F) for 1 h, oil quench
Temper 400 °C (750 ° F) for 1.5 h, oil quench	Temper 593 °C (1100 ° F) for 1.25 h, oil quench

The variation in energies for both lots was low: One lot had a coefficient of variation of 0.04 (an acceptable variation for impact verification specimen),
and the other lot had a coefficient of variation of 0.02 (a very low variation). However, to attain results of this quality for production lots of approximately1200 specimens, extremely well controlled processing is necessary, and typically double tempering, stress relief, cryo-treatment, and other steps are used to fine-tune the process for a given heat treating shop.

It is our experience that the specifics of the heat treatment should not be dictated to the shop. Each heat-treatment shop is different and needs leeway to adjust the process to best suit the equipment. Currently we use three shops and each uses a different process. All are capable of attaining similar quality specimens (after climbing difficult learning curves). A typical quality for impact verification specimens is characterized by a coefficient of variation (the ratio of the standard deviation to the average absorbed energy) of less than 0.04. The highest quality specimens approach coefficients of variation near 0.02.

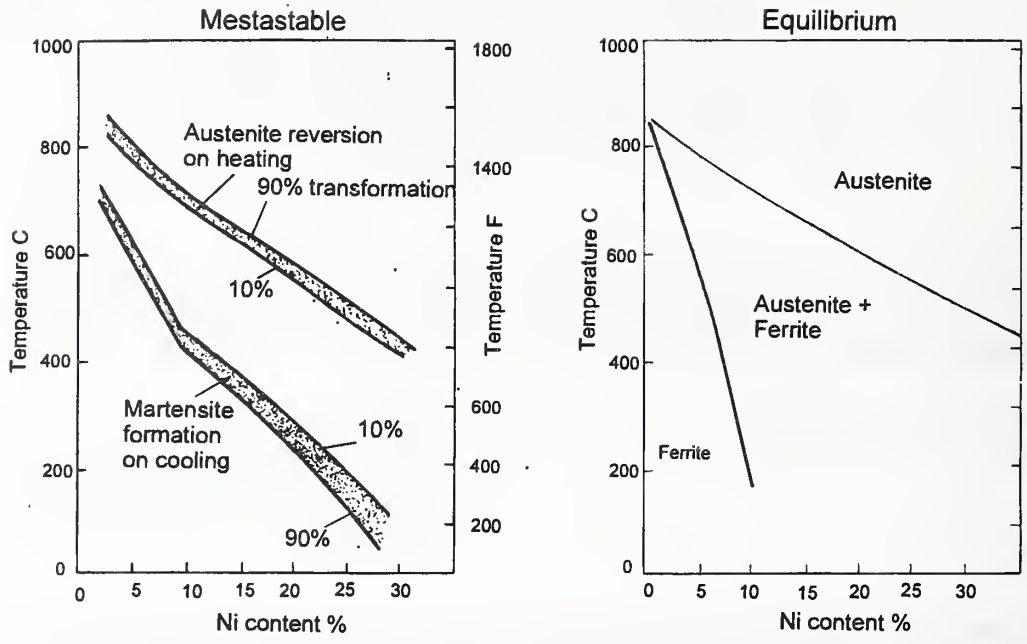


Figure 3: Phase diagrams for the Fe- Ni system.

5.1.2 **Type T-200 steel**

The T-200 steel is an 18 Ni, cobalt-strengthened maraging steel. This alloy can be solution-treated at 900 to 925 °C, control-cooled, grain-refined using multiple heating and cooling cycles near 760 to 815 °C, then aged to attain the appropriate strength/toughness combination. We age to produce a low-strength, high-toughness material. Recommended aging for a hardness of 30 HRC is 315 °C for 6 h. In general, the aging reactions are more sensitive to temperature than time.

The phase transformations for the T-200 steel that are of most interest are the martensite transformation on cooling, and the formation of austenite on heating (holding at temperature). As shown in **Figure 3**, the martensite in 18 Ni alloys is quite stable during heating to temperatures approaching 540 °C (1000 °F), which makes the aging of the martensite possible. However, substantial amounts of reverted austenite can form in Co-free maraging steels (and in other maraging steels) during aging treatments at temperatures of less than 540 °C (1000 °F), and it is not clear whether reverted or retained austenite would adversely affect the scatter in the Charpy impact energy.

For our alloy, we find that annealing temperatures of about 815 to 870 °C (1500 and 1600 °F) are high enough to avoid the two-phase region and produce a fully annealed structure, and low enough to avoid significant grain growth.

Most research does not include aging data for temperatures as low as 315 °C (600 °F), because it is not of commercial interest. There has been some indication, however, that different precipitates are formed when the alloys are aged at low temperatures. A study on an 18 Ni Co-containing 350 grade maraging steel showed distinct differences in the precipitates formed above and below 450 °C (845 °F). Ni_3Ti precipitates are formed in T-200 alloys at high aging temperatures, but actual precipitation probably doesn't occur at low aging temperatures (315 °C for 3 h). It is likely that clusters of Ni and Ti atoms cause the strengthening at low aging temperatures, and the toughness is lower for these under-aged clusters in maraging steels than it is for peak-aged steels (apparently because clusters or coherent precipitates restrict cross-slip in the matrix and Ni_3Ti precipitates allow more homogeneous slip).

Maraging steel can become embrittled during high-temperature solution treatments. The embrittlement is caused by precipitation of Ti(C,N) at grain boundaries during cooling, and can be retained even following re-annealing. Quenching from high temperature prevents the precipitation and subsequent embrittlement.

We have also found that quenching from high temperatures results in higher toughness (lower hardness) for our alloy, as indicated by the data in **Figure 4**. Quenching from the annealing temperature clearly results in a softer material, and the difference between the hardness of the air- cooled and water-quenched material is retained after aging. We found a difference of about

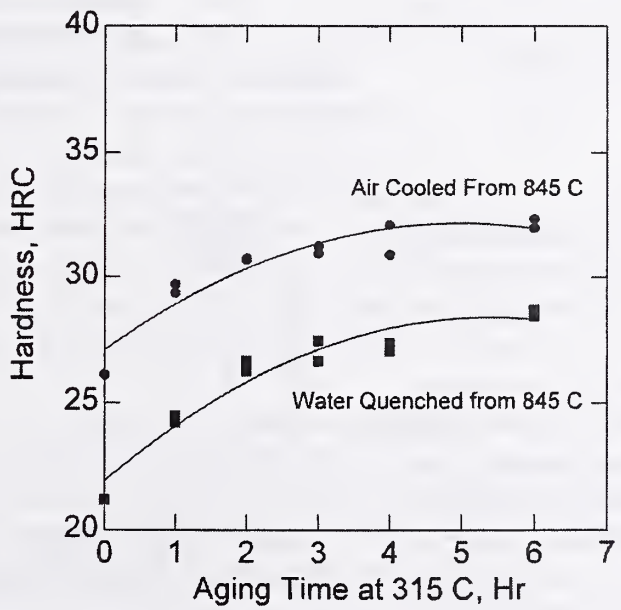


Figure 4. Hardness for air-cooled and water-quenched samples.

5 HRC, which is expected to result in a significant increase in the toughness of the material. Initial heat treatments on the new heat of T-200 material provide a general understanding of the energy levels that might be expected from the material. The mechanical test results for various heat treatments are shown in **Figures 5 through 7**. In **Figures 5 and 6**, the samples were annealed at 954 °C (1750 °F) for one hour and air-cooled, then re-annealed at 760 °C (1400 °F) for 1 h and air-cooled. These samples were then divided into five groups and aged at 260, 290, 315, 345, and 370 °C (500, 550, 600, 650, and 700 °F) for three hours and air cooled. The data show the relationship between the impact toughness and the hardness of the material for these heat treatment conditions. The data in **Figure 7** are similar to those in **Figure 5**, but these samples were annealed at 900 °C (1650 °F) for 1 h and water- quenched, then reheated twice to 675 °C (1250 °F) and water quenched as a grain refinement treatment, and re-annealed at 815 °C (1500 °F) for 1 h and air-cooled prior to aging at 315 and 370 °C (600 and 700 °F) for 3 h. Other variations of these two heat treatment schedules produced similar results. Overall, it appears that this T-200 material can be aged to produce Charpy specimen having impact energies of near 215 J (160 ft·lbf).

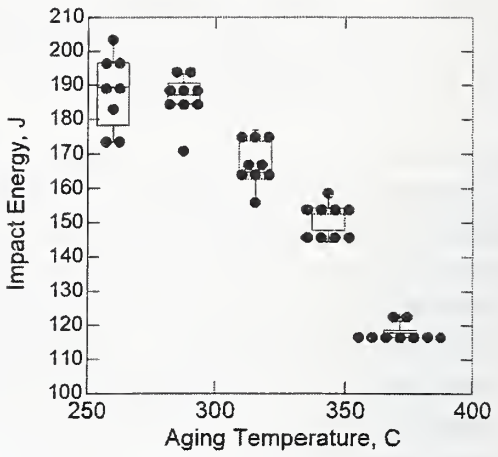


Figure 5. Impact energy of samples annealed at 954 °C and 760 °C, then aged.

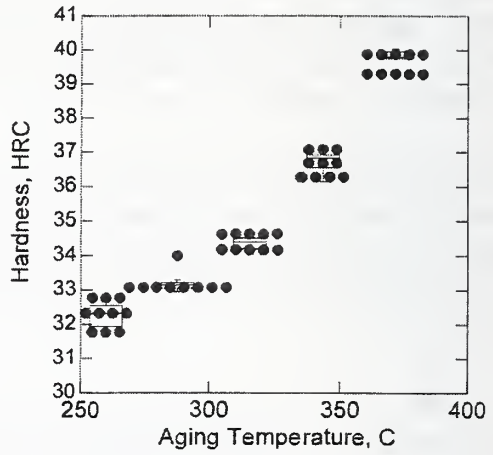


Figure 6. Hardness of samples plotted in figure 5.

5.2 Sampling

A production lot of approximately 1200 specimens are heat treated together as a single furnace load. A spatial (not random) sample of at least 100 specimens is removed from the heat treating baskets for pilot-lot evaluations. As shown in **Figure 8**, a spatial sample allows us to evaluate and minimize any correlation between the variation in energy of the samples to their position in the heat treatment baskets. If the pilot-lot sample is acceptable, the remaining specimens in the production lot are completed. An additional 30 random samples are removed from the production lot (following delivery to NIST). These

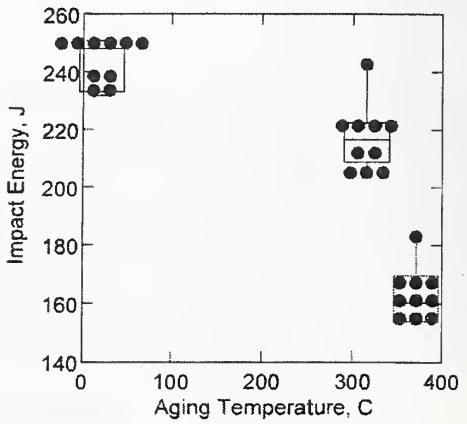


Figure 7. Impact energy of samples annealed at 900 °C, 675 °C and 815 °C prior to aging.

specimens are either used for evaluations of the production lot or held as control samples for future testing.

5.3 Machining

5.3.1 Process

Prior to heat treating, the square bars are cut to approximately 56 mm long blanks and ground to finished length. Then one end of the specimen blank is stamped with 'NIST', and other with a series number and a serial number. The series number identifies the production lot and the energy level (LL for low energy, HH for high energy, and SH for super high energy). The serial numbers range between one and the total number of specimens in the production lot. For the specimens made with 4340 steel, the surfaces are all ground to nominal size to remove surface flaws that might result in quench cracking during the heat- treatment operations.

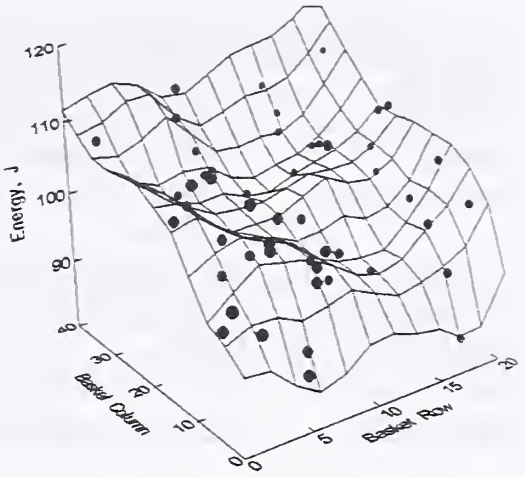


Figure 8. 4340 data showing the primary variation in the specimens correlates to the position of the specimen in the heat-treatment basket.

From the production lot of heat-treated specimen blanks, 100 are machined to final dimensions for pilot lot testing (**Figure 9**).

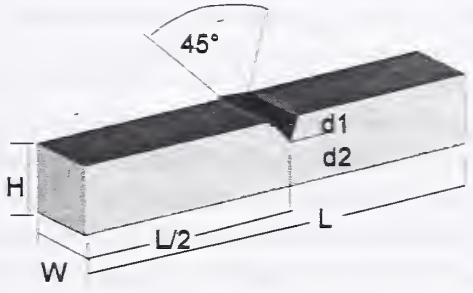


Figure 9. A Charpy V-notch sample with dimensions labeled in reference to Table 6.

5.3.2 Machining requirements

The dimensional requirements for NIST verification specimens, given in **Table 6**, meet or exceed the ASTM E 23 specifications. This minimizes variations in impact energy due to physical variations in the specimens. Also, the notch centering and the length tolerance for NIST specimens are equivalent to the ISO Standard 164, which permits the specimens to be used in impact machines with end-centering devices. The NIST requirement for surface finish is also equivalent to the ISO 164 requirement. All of these dimensional requirements can be met with standard machining practices.

Specimen notches are form ground on a surface grinder (machining with a fly cutter or multi-tooth cutter is not permitted). To avoid "burning" or cold working the material at the base of the notch, the next to the last cut is required to remove more than 0.25 mm and less than 0.38 mm and the final cut must not remove more than 0.12 mm. When the specimens are finished and ready for shipment, they are given a protective coating of oil.

Table 6. Dimensional requirements for NIST Charpy impact verification specimens.

Height (H)	10 mm, ±0.03 mm, with adjacent sides square within 90° ±9 min
Width (W)	10 mm, ±0.03 mm,
Length (L)	55 mm, +0.00 mm, −0.3 mm
Notch position L/2	27.5 mm ±0.2 mm, perpendicular to the longitudinal axis of specimen within 90° ±9min
Notch radius	0.25 mm, ±0.025 mm, with radius tangent to the notch angle
Notch depth (d1)	2 mm, ±0.025 mm
Notch angle,	45° ±1°
Ligament depth (d2)	8.0 mm, ±0.025 mm
Surface finish	1.6 µm on notched surface and opposite face; 3.2 µm on other surfaces

5.4 **Hardness Testing**

5.4.1 **Process**

Two hardness measurements are made on each of the pilot-lot samples, at positions approximately 10 mm from the specimen ends on the face opposite the notch. The two measurements are averaged to estimate the hardness of the sample.

The hardness criteria for verification specimens relate to three practical aspects of the impact test: (1) The minimum hardness requirement for low-energy lots assures an appropriate impulse load is transferred to the machine frame on impact to verify adequate mounting and overall stiffness of the machine. (2) The minimum hardness requirement for low-energy lots also determines the direction in which the 4340 impact specimens exit the machine.(3) The specimen-to-specimen variation in hardness provides an indication of the variation in energy of the specimens (particularly for the higher-energy specimens).

The verification specimens are produced so that different energy ranges leave the machine in different directions. Specimens with hardness of greater than 44 HRC leave the machine in a direction opposite to the direction of the swing, and are needed to evaluate how well the shrouds on U-type impact machines are functioning. Specimens with a hardness less than 44 HRC typically exit the machine in the same direction as the swing of the pendulum.

In practice we find that when the variation in hardness exceeds ±0.5 HRC the quality of the lot is questionable (i.e., the variation in energy is likely unacceptable). As shown in **Figure 10**, the correlation between energy and hardness is much more useful for evaluating variations

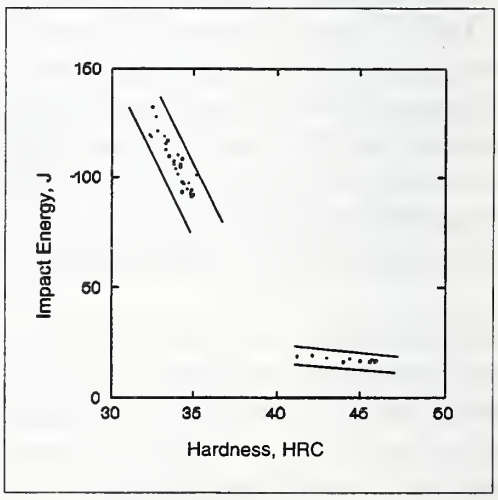


Figure 10. Hardness data for low and high energy verification specimens.

of high-energy (lower hardness) specimens, because at high hardness, the slope of the trend decreases significantly. So, although hardness evaluations have worked well as a quality-control procedure in our program, hardness data are principally used to estimate the impact toughness of the specimens and to assure that the low-energy specimens exit the machine in the required direction.

5.4.2 **Requirements**

The average hardness of the pilot lot samples must be within ±1 HRC of the targeted hardness, unless otherwise agreed.[2] This requirement is most important for the low-energy specimens, which normally need to have a hardness of 44 HRC or more to exit the impact machine properly.

5.5 **Impact Testing**

5.5.1 **Process**

If the dimensional measurements of the specimens and the hardness results are acceptable, the pilot-lot specimens are divided into three groups of 25 (one group is tested on each of the three master impact machines) and the extra 25 specimens are held in reserve for any additional testing that may be required. In dividing the specimens into groups, the furnace locations from which the specimens were taken are considered and the groups are balanced accordingly.

The certified energy value for a production lot of verification specimens is defined as the grand average of the 75 specimens tested (25 specimens on each of the 3 master machines). In addition to the grand average impact energy, the standard deviation, sample size, and several other statistics are calculated for the verification set. These statistics are used to determine the acceptability of the lot and the performance of the machines. All 75 specimens are included in these calculations, with the following three exceptions: (1) specimens that are determined to be outliers as defined in section 6.5.2, (2) specimens having the same lateral expansion but significantly different energies, and (3) specimens with flaws apparent on their fracture surfaces.

5.5.2 **General statistics**

The average energy and grand average energy are defined as

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \ldots x_n}{n}, \qquad (1)$$

where n is equal to 25 for calculating the averages of each machine, and n is equal to 75 for calculating the grand average for the pilot lot.

---

[2] When using a single impact machine, the same amount of impact energy may be indicated by materials having different yield strengths. These same materials tested in another machine may indicate different values of impact energy. The difference is usually greater for the stronger materials, presumably due to the faster rate at which peak loading occurs. To accentuate these differences, materials of high yield strength are specified for the verification specimens at each energy level. These requirements are normally monitored by making hardness assessments rather than tensile testing.

The standard deviation is defined as

$$s = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots (x_n - \bar{x})^2}{n-1}} \tag{2}$$

The pooled standard deviation is defined as

$$s_p = \sqrt{\frac{s_1^2 + s_2^2 + s_3^2}{P}}\,, \tag{3}$$

where subscripts 1, 2, and 3 indicate the standard deviation of the 25 samples tested on the three master machines, and P is equal to three.[3]

The sample size, which represents the minimum number of specimens from a given production lot that should be tested in a verification test, is defined as

$$n = \left\lceil \frac{3s_p}{E} \right\rceil^2\,, \tag{4}$$

where E is 1.4 J or 5 % of the mean energy, whichever is greater. For example, for the low-energy specimens E is equal to 1.4 J, so the maximum pooled standard deviation allowed for a sample size of 5 is

$$s_p = E\frac{\sqrt{n}}{3} = 1.4\frac{\sqrt{5}}{3} = 1.04\,J, \tag{5}$$

which indicates a CV of around 0.07 for specimens with an average energy of 16 J (1.04/16 = 0.07).

For the higher-energy specimens, E is taken as 5 % of the average energy for the lot and the maximum pooled standard deviation allowed for a sample size of 5 is

$$s_P = (0.037)\ (\text{average energy})\,, \tag{6}$$

and in this case the CV is 0.037 by definition (s_P /average energy).

---

[3] The choice of the value of standard deviation depends on whether all the machines used to determine the reference value met the requirements for variability, k (see equation 7). If all the machines met the requirements, the value of s shall be equal to the pooled standard deviation. If all the machines did not meet that requirement, s shall be equal to the largest of the standard deviations of the machines considered separately.

The outlier analysis is performed using box-and-whiskers plots to provide a graphical summary of the data and identify outliers.[4] Outliers are defined as values that are lower than the first quartile or higher than the third quartile by more than 1.5 times the absolute difference between the first and third quartiles. If a lot has more than 5 % outliers, it may be rejected.

The variation in energy values is calculated for each machine using a ratio of the standard deviation for the particular machine over the pooled standard deviation. This ratio, k, is expressed algebraically as

$$k_1 = \frac{s_1}{s_p}, \quad k_2 = \frac{s_2}{s_p}, \quad k_3 = \frac{s_3}{s_p}, \tag{7}$$

here s_n and s_p are the individual and pooled standard deviations for the three machines respectively. If the k ratio of any of the three machines exceeds 1.25 (assuming 25 specimens tested per machine), the variability in energy values due to that machine is questioned and appropriate actions are taken (repairs to the machine, testing of additional samples, etc.).[5]

5.5.3 **Uncertainty calculation**

The uncertainty of a single specimen in a given lot can be determined by combining three components of uncertainty: within-machine uncertainty (s_P), uncertainty due to machine bias (s_B), and the uncertainty of specimen homogeneity (s_H). The total uncertainty is given by

$$u_T = \sqrt{\left(s_P^2 + s_B^2 + s_H^2\right)}. \tag{8}$$

The within-machine uncertainty is the "pooled" standard deviation (see eq (3)) based on 25 verification specimens tested on each of the 3 master machines. The degrees of freedom associated with s_P is 72 (i.e., 25 + 25 + 25 - 3).

The uncertainty due to machine bias accounts for possible bias in the observed averages associated with each master machine. The value of s_B can be quantified using a technique called "BOB" which models the unknown biases with a Type B uncertainty distribution.

The final component of uncertainty, s_H, can be thought of as a correction for specimen inhomogeneity and is typically based on engineering judgment. It is common practice to set the

---

[4] An outlier is defined statistically, but a specimen identified as an outlier is not removed from the analysis unless it shows physical evidence of jamming, material flaws, or other reasons for atypical behavior.

[5] If the k ratio of a machine is greater than 1.25, the results of this machine can be questioned by the contractor who supplied the specimens. This is considered a basis for retesting another group of 25 specimens prior to determining the acceptability of the lot.

number of degrees of freedom associated with a Type B component of uncertainty, such as s_H, equal to infinity.

5.5.4 **Energy requirements**

The most important requirement is the variability in impact energy of the specimens. Our contracts allow us to reject a lot with a sample size of more than 5.

A lot can also be rejected if the average energy is outside the range specified in Table 6. The certified energy of the specimens must fall within the ranges of 14 to 20 J (10 to 15 ft·lbf) for the low energy level, 88 to 136 J (65 to 100 ft·lbf) for the high energy level, and 176 to 244 J (130 to 180 ft·lbf) for the super-high energy level, unless otherwise agreed.

6. **Certification and Acceptance of a Pilot Lot**

6.1 **Process**

Acceptance of a new batch of verification specimens is based on the data obtained from the pilot lot of 100 specimens, taken from a heat-treatment batch of approximately 1200 specimens. Although impact energy is the most important criterion, other criteria are also evaluated to determine the consistency and quality of the verification specimens. The pilot lot data (impact energy, hardness, and dimensional measurements are processed using a computer program to provide standardized output for review in determining the acceptability of the lot.

If the data indicate that the pilot lot is acceptable, the contractor is advised to machine the remainder of the production lot and submit it for final acceptance. If the random samples removed by NIST from the production lot are acceptable, a certified energy value is assigned to the lot and it is placed in inventory.

The certified energy of the lot is defined as the grand average energy of the lot. The number of specimens in a verification set is determined by the sample size calculation. Typically, a set size of five is used (for lots with sample sizes of three, four, or five), but occasionally sets having more or less than five samples are distributed.[6]

The number of degrees of freedom associated with each of the three components of uncertainty can be combined using the Welch-Satterthwaite formula to obtain the effective degrees of freedom associated with the total uncertainty, u_T. The effective degrees of freedom are used to determine the appropriate coverage factor for the confidence intervals.

6.2 **Requirements**

A pilot lot can be rejected for use as verification specimens if:

1. The verification specimens do not meet the dimensional requirements given in section 6.3.

---

[6] We routinely reject lots with sample sizes greater than five, but when stocks are very low we have occasionally accepted lots with larger sample sizes.

2. The hardness and energy levels are not within the ranges specified in Table 7.[7]

Table 7: Required ranges for verification specimens.

Energy level	Low	High	Super-high
Absorbed Energy (J)	14 to 20	88 to 136	176 to 244
Hardness (HRC)	>44	∓1 of avg	∓1 of avg

3. The sample size for the lot exceeds 5.
4. The number of outliers exceeds 5 % of the number of specimens impact tested in the pilot lot (4 is the maximum for a pilot lot of 75).
5. The difference between a machine average and the grand average is greater than the larger of 1.4 J or 5 % of the grand average.[8]
6. The results from one of the three machines show excessive variability, according to the k ratio.[9]
7. The microstructure of the low energy and high energy specimens is not 100 % martensite (no ferrite, austenite, or bainite should be visible).

6.3 **Reports**

The pilot-lot data calculations are done by a computer program to provide a consistent appearance and quality for our records. The evaluation report documents the lot identification, the reference machine it was tested on, and the energy and hardness of each specimen that was tested. The calculated values in the report are as follows: (1) the grand average energy and standard deviation, (2) the average energy and standard deviation for each machine, (3) the average hardness and standard deviation, (4) the pooled standard deviation in energy for the lot, (5) the sample size, and (6) the k ratio of each machine. A set of standardized plots in the report shows outlier data, the distribution in energy for each machine, and the combined distribution in energy for the three reference machines. The data collected from dimensional measurements on the specimens are kept separately. The evaluation report and raw data for the pilot lots are filed for future reference and a copy of the report is sent to the contractor who supplied the pilot lot (with our comments).

[7] The average hardness values of the high and super-high energy specimens are not specified, but we require a maximum variation of ±1 HRC for the lot.

[8] In this case, both the quality of the specimens and the performance of the machine are questioned. Appropriate actions are taken that are agreed to between the NIST and the contractor supplying the specimens.

[9] In this case, both the quality of the specimens and the performance of the machine are questioned. Appropriate actions are taken that are agreed to between NIST and the contractor supplying the specimens.

7. Distribution

7.1 Packaging

Specimens are drawn from the production lot at random to make up the sets of impact verification specimens. These sets are distributed for verification testing. Each set of verification specimens is retained henceforth in the sets, as originally drawn.

If a purchaser can demonstrate that one or more specimens of a set are defective, the set is replaced without charge.

7.2 Information

Each set of verification specimens is accompanied by a certificate (Appendix 1) that gives the following information: name, address, and telephone number of NIST contacts; the test temperature; the identification of steel used; the designation number of the practice or practices whose specifications for verification specimens are met by the specimens supplied.

8. Customer Certification Procedure

8.1 Process

The results of a verification test are returned to NIST, along with the broken specimens and a questionnaire that is filled out by the customer. Information from each of these three sources is used in the evaluation of the test. Based on the results of this evaluation a letter is written to the customer. If the results are acceptable, a verification letter and accompanying verification sticker serve as documentation that the machine meets the requirements of ASTM Standard E23. If the results are unacceptable, the letter explains why we think the test does not meet the requirements of E 23 and suggests how the machine might be brought into compliance.[10]

8.2 Customer Questionnaire

The information provided by the customer is used to help us understand anomalies in the test data and provide background that allows us to better advise the customer. If test results are uniformly high, for example, the questionnaire might be referenced to determine how the test temperature was measured and the last time the temperature equipment was calibrated, which might explain the result (test conducted at wrong temperature). Other, nontechnical information is also provided by the questionnaire that is used to update our database. A copy of the questionnaire is given in Appendix 2.

8.3 Test Data

The verification test results are calculated and compared with the certified value of the lot, and with the results of previous verification test results for the machine. A machine is classified as unacceptable if the difference between the average energy of the machine being verified and the

---

[10] Customers are not required by E 23 to have their machines verified by NIST. They can verify the results of the test themselves, or have a private testing company verify the test results. ASTM E 23 does require that the energy value of the impact verification specimens be "established on the three reference machines owned, maintained, and operated by NIST in Boulder, CO".

certified value is greater than 1.4 J or 5 % (whichever is larger) of the certified energy of the verification lot.

8.4 **Examination of Broken Specimens**

The specimens are checked to determine the following information: (1) if the anvil marks indicate that the specimens were centered for the test, (2) if the striker mark indicates that the striker on the impact machine was centered, (3) if the anvil markings indicate excessive or unusual wear, (4) if the size of the shear lips on the specimens indicate that the test was done at the proper temperature, and (5) if markings on the fracture surfaces of the specimens show material flaws or unusual textures.

These observations are used either to remove a specimen from the data analysis (in the case of a flaw or off-center strike), or as a basis to fail the test due to worn anvils, etc. More specific information on how and why the specimen are examined in the NIST procedure are given in Special Publication 960-4.

8.5 **Customer Letter**

Based on our judgement and the requirements of ASTM E 23, a pass or fail letter is developed for the customer. The letters are composed using a word processing program that is integrated with our database, and with a list of standard paragraphs covering commonly observed problems with verification test results. The program merges customer data with the selected standard paragraphs, and then allows final editing for the addition of more specific comments, if applicable. The letter also includes a table that presents the customers' data and the values that were computed by the program to evaluate the data.

If a customer fails the verification test, he/she is typically contacted by fax, phone, or email to discuss the results. If a customer passes the verification test, the letter serves as a file record.

8.6 **Verification Sticker**

A verification sticker (to put on the impact test machine) is mailed with each pass letter. The stickers have a NIST logo and give the serial number of the machine, the date of the next verification, and the range in energy over which the machine is verified.

The stickers are made using a Brady 200M label printer and Codesoft version six software by Teklynx.

The inclusion of stickers in the customer letters was initiated in September 2001.

9. **Program Controls**

9.1 **Impact Machines**

The impact machines are inspected and adjusted by NIST personnel, and experts contracted by NIST. Critical direct verification measurements were made when the machines were installed, and are made when a change in the performance of a machine is noted. Example data for the master machines are given in Appendix 3.

The performance of the impact machines is routinely evaluated for each lot of specimens tested. This evaluation is principally a comparison of the mean and standard deviation of each machine to the other machines used in the program. The performance of the machines are compared as each pilot lot is tested, and these results are compared with the past performance of the machines.[11] A plot showing the average energy of each machine and the grand average for each pilot lot is updated for each pilot lot tested, to document and evaluate the relative performance of the impact machines. Example data are given in Appendix 3.

A log book on the machines is maintained that contains records for the "daily check" procedures that are conducted on the machines prior to testing a pilot lot: these records allow us to track the friction and windage, and other factors that affect the performance of impact machines. The log book also documents maintenance to the machines and the number and types of specimens tested.

A reserve of impact verification specimens (from past pilot lot tests) are kept and serve as control specimens. When a change to a machine is suspected, due to its relative performance, a set of control specimens can be tested and compared to the original performance for this machine with these specimens. Control specimens are also used to check machines following a repair.

9.2 **Measurement Equipment Used in the Verification Program**

A Newage Deltronic hardness tester is used to measure hardness. The hardness tester is calibrated annually by Leco Corporation. The hardness tester is checked with calibration blocks prior to each use. An optical comparator is used to measure the notch angle, notch depth, notch radius, and L/2 (notch centering in relation to specimen length). The optical comparator is a Deltronic Model DH 216 and is equipped with an MPC-5 readout. The comparator is calibrated annually by Precision Gage, Inc. Both the hardness tester and the optical comparator read directly to a personal computer using NIST developed software.

Mitutoyo, Model CD-6"C, digital calipers are used to measure specimen length, width, and thickness. The calipers are calibrated annually by Precision Gage, Inc. The calipers are checked with a one-inch calibration block prior to each use. The caliper data are automatically stored on a personal computer.

Squareness is measured with a gage manufactured by Laboratory Testing, Inc. The gage was manufactured using the drawing in ASTM Standard E 23. The gage is calibrated annually by Laboratory Testing, Inc. The gage is checked with a calibration block furnished by, and annually calibrated by, Laboratory Testing, Inc. All calibrations by outside companies are traceable to NIST.

---

[11] The impact machines have characteristic differences from one another in energy level and variation. Changes in these relative differences indicate changes to our program, and are investigated to determine the cause.

9.3 **Specimens**

The quality and consistency of the verification specimens is first controlled by the steel used for their production. Our contractors are shipped bundles of steel bar that are coded with reference to ingot location, and production lots are made using steel from a given bundle. This is our best assurance that the steel used for a given production lot is as similar as possible. In the event that some portion of the bar contains melting or rolling flaws, this procedure would help us to more quickly identify and remove this material from the stock.

Our second control of specimen quality is careful sampling and pilot lot evaluations. In our experience we have found that geometric rather than random sampling produces a better estimate of the mean energy for our pilot lots. Our samples are taken from predetermined positions within the heat-treating baskets and labeled.

Our final control involves a feedback loop using data from customer verification tests. As customer data are collected they are stored in a database, and pass/fail ratios can easily be calculated for a lot of verification specimens that is questioned by either a customer or ourselves. If these data show normal ratios, this is strong evidence the average energy of the lot was accurately estimated by our pilot-lot sample. If these data show more machines than normal are failing using a particular lot of specimens, and the mean energy of the customer data is significantly different from the certified energy value of the lot, this is evidence that the certified energy value of the lot has changed or that the average energy determined for the lot was not an accurate estimate.[12]

9.4 **Customer Evaluation and Service**

In an attempt to control the quality of our customer assessment, we look at both current and past tests results for the machine. This often helps in understanding a customers' problems and allows us to better help the customer with comments concerning the performance of the machine. To provide prompt service for the program, we have two back-up personnel who are capable of filling in to cover the day-to day operation of the program.

To preserve good documentation of customer verification tests, we save the test specimens for one year. We save customer letters for two years as a hard copy. Digital files of letters are kept indefinitely, and all database information is saved.

9.5 **Database**

The software used in the program is managed by one person only. No changes are made without adequate follow-up. The personnel using the software are always consulted before and after making software revisions. The database (and software) is backed-up to tape on a regular basis.

---

[12] In the last 10 years approximately 3 lots have been suspected of having inaccurate certified energies assigned to them. In all 3 cases, sampling is suspected to have caused the error. The stability of the specimens has not been suspected as the cause (because the energies increased rather than decreased).

10. Support of National and International Impact Standards

The data gathered from customer verification tests and from our pilot lot tests provide a unique source of information for statistical studies on impact testing. These data allow us to review how specific designs, capacities, and ages of impact machines perform under the rules of the various impact standards used in the world. They also allow us to evaluate how well our estimates for the average energies of pilot lots compare with the average of all the machines in the world that tested them.

The principal use for these data is to help address issues concerning the indirect verification rules of ASTM and ISO impact testing standards. For example, currently there are major differences between the pass/fail range for ASTM and ISO verification tests, and our data can be used to make a strong argument that the ISO tolerance is too large (and the ASTM tolerance may be too small). Since the data are from actual tests, and include results from many countries, we can more accurately evaluate (and demonstrate) the impact of verification rules to our customers and the standards community than can any other country in the world. Our database is currently estimated to have results for more than 14000 tests (sets) and approximately 150 pilot lots.

11. Education and Training

11.1 Customers

Misunderstandings with our customers are minimized once they understand the NIST program, and how they can best take advantage of it. To help educate our customers we have developed a video that provides an overview of the Charpy test and the NIST verification program. We also have a brochure detailing our specimen evaluation procedures, Special Publication 960-4. Occasionally we have provided group training to companies that manufacture and repair Charpy impact machines.

11.2 Staff

General experience is the most valuable education and training for our staff, particularly experience gained from talking to customers. In addition, we attend ISO and ASTM meetings to gather feedback on the program. Comments at these meetings, and other technical meetings concerning impact testing, help to keep the program and personnel on track.

12. Safety Considerations

The operation of impact machines requires good safety practices to avoid injury. During our impact tests, the laboratory is locked and signs are posted on the doors to indicate that testing is in progress. Partitions are used to shield the operator and other equipment in the room from flying specimen halves as they exit the machine after impact. Safety glasses are worn when impact tests are conducted.

Appendix 1. Sample Charpy Verification Certificate

Example of a Certificate that was distributed with Charpy verification samples in 2000.

National Institute of Standards & Technology

Certificate

Standard Reference Materials ®

2092 - Low-Energy
2096 - High-Energy
2098 - Super High-Energy

Verification Specimens for Charpy V-Notch Impact Machines

Lot No.:

Standard Reference Materials (SRMs) 2092, 2096, and 2098 are intended primarily for the verification of Charpy V-Notch machines in accordance with the current ASTM Standard E 23 [1]. Each SRM consists of a set of individual 10 mm × 10 mm × 55 mm specimens needed to perform one verification. These SRMs comply with both ASTM Standard E 23 and International Organization for Standardization ISO/DIS 12736 dimensional requirements [2].

Material Description: SRMs 2092 and 2096 are made from 4340 alloy steel. SRM 2098 is made from a high strengthmaraging steel. The bars are finished to length, stamped, heat-treated, and machined in SRM specimen lots ofapproximately 1200. Each specimen has a lot number and an identification number (three or four digits) stamped on oneend of the specimen. Additional information can be found in References [3-5].

SRM Certification Procedure: Specimens taken at random from each SRM lot are tested by the NIST Materials Reliability Division on Charpy V-Notch reference machines. The specimen data generated are then statistically evaluated to assure the homogeneity of the lot, establish the certified value, and determine the number of SRM specimens required for a user to perform a valid test. See Table 1 for a list of the approximate energy ranges within which the individual certified values should fall.

If certified values are required immediately after testing, contact the NIST Charpy Program Coordinator as follows: telephone (303) 497-3351; fax (303) 497-5939; or e-mail vigliotti@boulder.nist.gov. The lot number and energy results of the tested specimens must be provided in order to obtain certified values by telephone or fax.

Expiration of Verification: The verification report issued on an acceptable machine is valid for one year from the date that the SRM was tested. If a user's machine is moved or undergoes any major repairs or adjustments, the current verification will be invalidated and the machine must be retested and reverified. The overall direction and coordination of the technical measurements leading to verification of test specimens and machines, evaluation of test results, and issuance of the report on machine conformance are under the direction of the NIST Materials Reliability Division, Boulder, CO.

The support aspects involved in the original preparation, certification, and issuance of these SRMs were coordinated through the NIST Standard Reference Materials Program by R.J. Gettings. Revision of this certificate was coordinated through the NIST Standard Reference Materials Program by C.R. Beauchamp.

Fred R. Fickett, Chief
Materials Reliability Division
Gaithersburg, MD 20899 Nancy M. Trahey, ChiefCertificate Issue Date: 14 May 2001 Standard Reference Materials Program

NOTE: THESE ARE NOT CERTIFIED VALUES. THESE ARE THE APPROXIMATE RANGES FOR EACH ENERGY LEVEL.

Table 1. Approximate Charpy SRM Energy Ranges.

SRM	No.	(J) (ft·lbf)
2092	13-20	10-15
2096	88-136	65-100
2098	176-244	130-180

Storage: The SRMs are composed of specimens anticipated to have an indefinite shelf life under normal storage conditions. Each specimen is coated with oil, wrapped in a corrosion inhibiting paper, and sealed in a plastic envelope.

It is recommended that the specimen be retained in this package to protect them from moisture until used. The protective oil coating should be wiped from each specimen just prior to testing.

Use: Prior to testing a Charpy V-Notch machine, the machine should be checked to assure compliance with the appropriate sections of the current ASTM Standard E 23 [1]. To comply with the testing procedures specified in the standard, SRM 2092 and SRM 2096 shall be tested at −40 °C ± 1 °C (-40 °F ± 2 °F). SRM 2098 shall be tested at 21 °C ± 1 °C (70 °F ± 2 °F). All SRM specimens are to be tested in accordance with the testing procedures of the appropriate sections of the current ASTM Standard E 23. All SRMs shall be tested at the same time. An acceptable machine will produce an average value within 1.4 J (1.0 ft·lbf) or 5 % of the certified energy value, whichever is greater, providing the specimens appear to have normal markings. Because the source(s) and magnitude of error for energy values at one energy level may not be the same at different energy levels, calibration or correction curves shall not be used.

Verification of User's Machine: The NIST Charpy Program Coordinator will issue a report of findings to the user's facility upon receipt of the fractured specimens and completed questionnaire. If the machine to be verified produces acceptable values and the specimens appear to have normal markings, this report will verify its conformance. If the machine produces values outside the allowable tolerance of the certified energy values or the specimens have abnormal markings, the report may suggest repair or replacement of machine parts, changes in testing techniques, or other appropriate corrective actions. Fractured specimens and completed questionnaires should be returned to the NIST Charpy Program Coordinator, Mail Code 853.07, 325 Broadway, Boulder, CO 80305-3328. A plastic, self-locking bag is provided for the return of broken specimens. The broken specimens shall be taped together as described in the wrapping instructions included with the questionnaire.

Important Information: Shipping charges for the return of broken specimens are the responsibility of the user. The mailing label provided with each SRM must be used to expedite shipping and, for overseas shipments, clearance by U.S. Customs.

Note to International Customers: Regular overseas shipments of broken specimens should be sent airmail so that after they are cleared by U.S. Customs, they can be forwarded directly to NIST-Boulder. If a more rapid shipping mode is necessary, choose an overnight delivery service that will handle U.S. Customs clearance **AND** will deliver directly to NIST-Boulder. Unless such delivery is assured, air freight packages may be returned to the customer by U.S. Customs.

REFERENCES

[1] ASTM E 23, Standard Test Methods for Notched Bar Impact Testing of Metallic Materials, Annual Book of ASTM Standards, **03.01**, ASTM, West Conshohocken, PA.
[2] ISO/DIS 12736, Metallic Materials - Impact Testing - Preparation and Characterization of Charpy V Reference Test Pieces for Verification of Pendulum Impact Testing Machines, ISO, Geneva, Switzerland.
[3] Siewert, T.A. and Schmieder, A.K., "Pendulum Impact Machines: Procedures and Specimens for Verification," ASTM STP 1248, ASTM, West Conshohocken, PA, (1995).
[4] Shepherd, D.A. and Siewert, T.A., "Interlaboratory Test Study for the Determination of Precision and Bias in Charpy V-Notch Impact Testing," ASTM Research Report E 28-1014, ASTM, Philadelphia, PA, (1991).
[5] Holt, J.M., "Charpy Impact Test - Factors and Variables," ASTM STP 1072, ASTM, Philadelphia, PA, (1990).
Users of this SRM should ensure that the certificate in their possession is current. This can be accomplished by contacting the SRM Program at: telephone (301) 975-6776; fax (301) 926-4751; e-mail srminfo@nist.gov; or via the
Internet http://www.nist.gov/srm.

Certificate Revision History: 14 May 2001 (updated email address for Boulder contact); 09 August 2000 (updated mail and zip codes for Boulder facility); 22 March 2000 (editorial revision); 26 July 99 (editorial revision); 20 February 97 (original certificate date).

Appendix 2: Sample Customer Questionnaire

Example of a customer questionnaire used for the Charpy program.

QUESTIONNAIRE FOR CHARPY IMPACT MACHINE VERIFICATION

IMPORTANT: This questionnaire contains information to help you perform a successful verification test. Energy results are required for verification. Other specific information is requested to help evaluate the condition of your machine. The questionnaire and the fractured specimens should be shipped to the Charpy Program Coordinator, NIST, Division 853, 325 Broadway, Boulder, CO 80305-3328. Phone: 303/497-3351 Fax: 303/497-5939.

Location of Machine

Company _____

Address _____

_____

City _____ State/ Province _____

Country _____ Zip/ Postal Code_____

Mailing Address for Verification Letter (if different from above)

Company _____

Address _____

_____

City _____ State/ Province _____

Country _____ Zip/ Postal Code_____

Test Machine (circle appropriate units where indicated)

1. Machine Manufacturer and Serial Number_____
2. What is the maximum energy capacity of the machine? _____ (J ft·lbf)
3. If the machine is adjustable, what capacity was used for this test? _____ (J ft·lbf)
4. The machine should be securely bolted to a concrete foundation or a steel block having a mass not less than 40 times that of the pendulum.

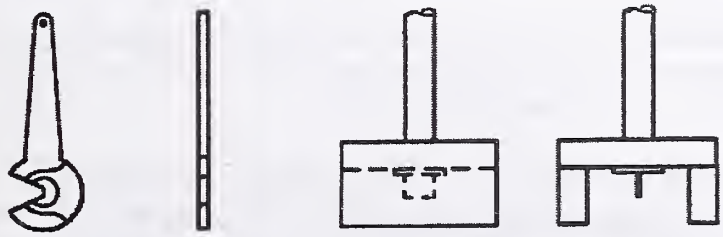
   (a) What type of bolts are used to mount the machine? (J, lag, etc.) _____

   (b) The machine should be level according to the current ASTM Standard E 23.

5. Is your machine equipped with a carbide striker? _____

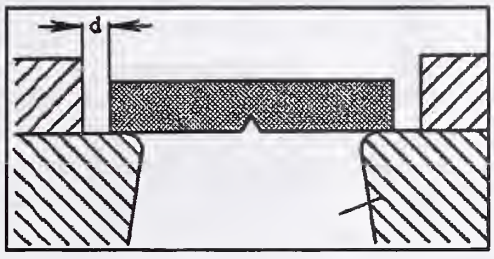
6. Is your machine equipped with carbide anvils? _____

7. Check the appropriate pendulum design below.



A _____ B_____ C (Other) _____

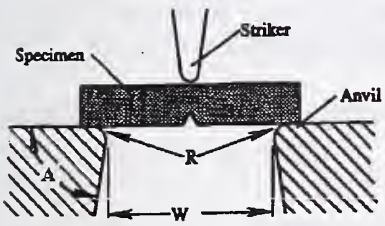
Side Support

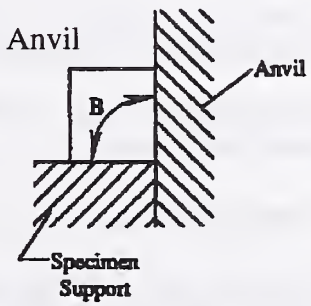


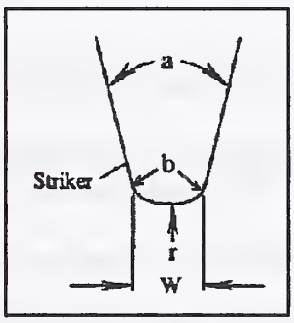
Please Sketch

8. If side supports or shrouds are used, what is dimension "d"? _____
Circle: (mm or in)

9. Your anvils and striker should conform to the dimensions below:

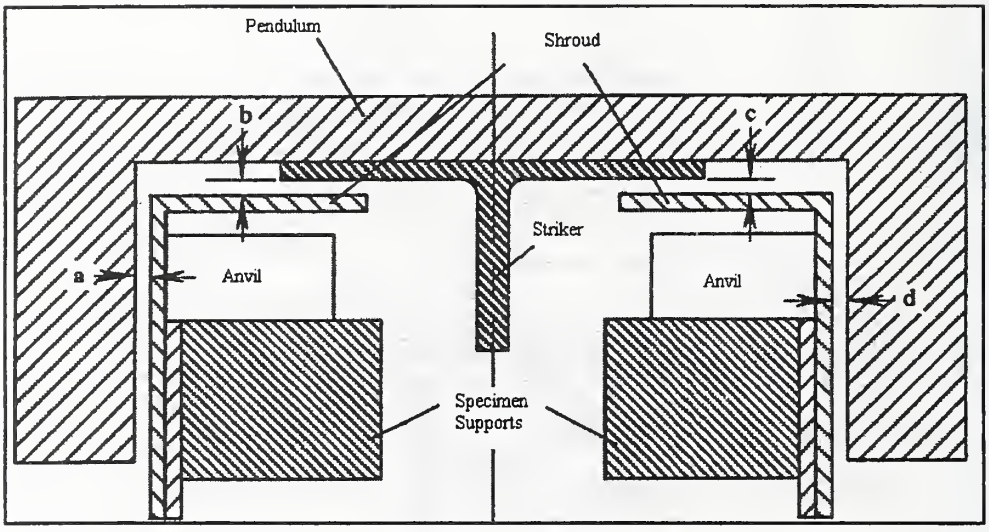






	Anvils		Striker
A:	80° approx.	a:	30° approx.
R:	1 ± 0.05 mm	r:	8 ± 0.25 mm
	(0.039 ± 0.002 in)		(0.315 ± 0.010 in)

W:	40 ± 0.05 mm	w:	4 mm approx.
	(1.574 ± 0.002 in)		(0.157 in)
B:	90° ± 10 min	b:	0.25 mm (0.010")



10. If shrouds are used to contain broken specimens, the following requirements should apply:
    (A) The shrouds should have a minimum hardness of 45 HRC.
    (B) The thickness of the shrouds should be approximately 1.5 mm (0.06 in).
    (C) Dimensions a, b, c, and d below should not exceed 1.5 mm (0.06 in).
    (D) If dimension "d" in item 8 is more than 13 mm (0.5 in), requirements (B) and (C) above do not apply.

11. The striker should pass through the center of the anvils within 0.40 mm (0.016 in).

12. With the pendulum in the free hanging position, engage the energy indicator. The indicator should read within 0.2% of the maximum energy range being used.

13. What is the friction /windage loss of your machine? _____
    (J ft·lbf)

    (a) Raise the pendulum to the latched position. Without a specimen in the machine, release the pendulum and permit it to swing 11 half cycles; after the pendulum starts its 11th half cycle, move the pointer to between 5 to 10 % of scale range capacity and record the dial reading.

    (b) Divide the value by 11, then divide by the maximum scale range of the machine and multiply by 100. The result, friction and windage loss, should not exceed 0.4 %.

14. With the specimen removed from the machine and the pendulum released from its latched position, what is the dial reading after one swing? _____
    (J ft·lbf)

This reading should be zero. If this reading is not zero and your machine is equipped with a compensated scale, please adjust the dial to read zero. If your machine is equipped with a non-compensated scale, please compensate the energy values for windage and friction by subtracting the windage and friction value calculated in item 13.

15. When was this machine last verified by the NIST? Date: _____

16. Is your machine equipped with a direct reading scale or a noncompensated scale? _____

IMPORTANT INFORMATION

To obtain accurate results the following procedures should be followed closely. For the NIST reference specimens the test temperature is near the ductile-brittle transition temperature of the steel. Therefore small variances in temperature and procedure may cause considerable error in energy values.

- The cooling bath should be placed directly beside the machine. This enables the operator to remove specimens from the bath and fracture them in the machine quickly.
- It is very important that the specimens be removed from the bath and fractured in less than 5 s. Taking longer than five seconds can increase the energy values, which may cause the low energy specimens to exceed the allowable energy limit.
- If your machine is equipped with a centering device, we do not recommend that you use it to center specimens when performing low temperature testing. Instead, we recommend the use of centering tongs as described in the current ASTM Standard E 23. The centering tongs should be cooled with the specimens.
- Verify temperature-measuring equipment at least twice annually. The measurement equipment can be checked immediately before the test by checking a medium with a constant temperature such as dry ice [−78.6 °C (−109.3° F)] or ice water [0.0 °C (32.0° F)].
- When testing super-high energy level specimens or other ductile materials, the anvils should be checked between each test for material left by the previous test.
- When the anvils are replaced it is recommended that practice specimens be broken before NIST specimens are tested.

TESTING TECHNIQUE

1. Test temperature for SRM 2092 low energy and SRM 2096 high energy level specimens should be −40 ± 1° C (−40 ± 2° F).

2. Test temperature for SRM 2098 super-high energy level specimens should be 21 ± 1° C (70 ± 2° F).

3. How long were the specimens held at temperature? (NIST recommends a minimum of 10 min)

_____

4. What instrument was used to remove the specimens from the bath and center them in the machine?

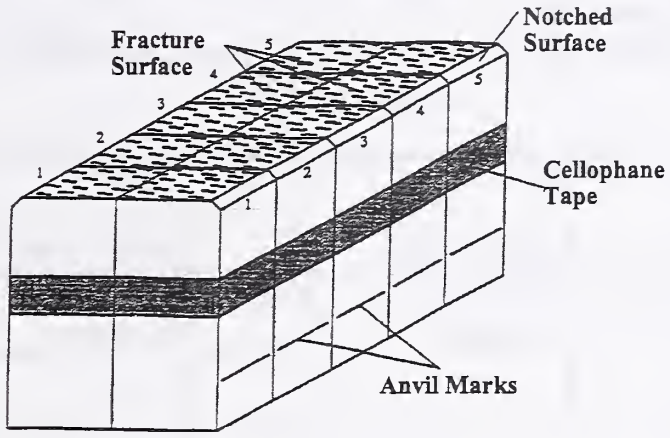
_____

STATE REASON FOR VERIFICATION

1. Compliance with annual ASTM Standard E 23 Indirect Verification _____
2. Changed striker and/or anvils _____
3. Moved machine _____
4. Changed bearings or pendulum _____

WRAPPING INSTRUCTIONS

To expedite the evaluation of your machine, please secure the 5 broken specimens (10 halves) from a particular energy series, as one unit with **clear cellophane tape** according to the following instructions. See diagram below.



1. Keep broken halves correctly paired (back to back) with the fracture surfaces facing upward and notched surfaces facing outward.

2. Coat the **FRACTURE SURFACES ONLY** with a light coat of oil. **DO NOT** use grease or coat in plastic.

3. Include this completed questionnaire with the fractured specimens.

4. Be sure that you use the **MAILING LABEL**, provided with the specimens, and attach the label so that it is clearly displayed on the **OUTSIDE** of the package. This will expedite delivery to the Charpy Coordinator. Customers returning specimens from outside the United States should include the following statement on the U.S. Customs Declaration:

   Contents include U.S. manufactured steel test bars being returned to the U.S. for evaluation and are valued at less than 10 U.S. dollars.

Sample Test Results Report

NOTE: Use ONE questionnaire only to report the NET ENERGY RESULTS of all energy levels used to test this machine at this time.

INDICATE ENERGY UNITS (circle units used)

Joules ft·lbf

Series _____ SRM 2092		Series _____ SRM 2096		Series _____ SRM 2098	
Specimen Number	Value	Specimen Number	Value	Specimen Number	Value
Average Value		Average Value		Average Value	

Date of Test _____
(Month/ Day/ Year)

Test Operator
PRINT_____ **Telephone**
_____

Test Operator
SIGNATURE_____ **FAX**
_____

Company Representative
PRINT_____ **Telephone**
_____

Company Representative
SIGNATURE_____ **FAX**
_____

If you require approval of your machine by the Defense Contract Management Command (DCMC), a DCMC representative should provide his or her signature and the DCMC seal to indicate that the preceding information was witnessed by a government representative.

_____

Print Name of DCMC Official **Seal**

_____

Signature of DCMC Official and Seal

_____

DCMC Office Location

Appendix 3. Machine Performance Data

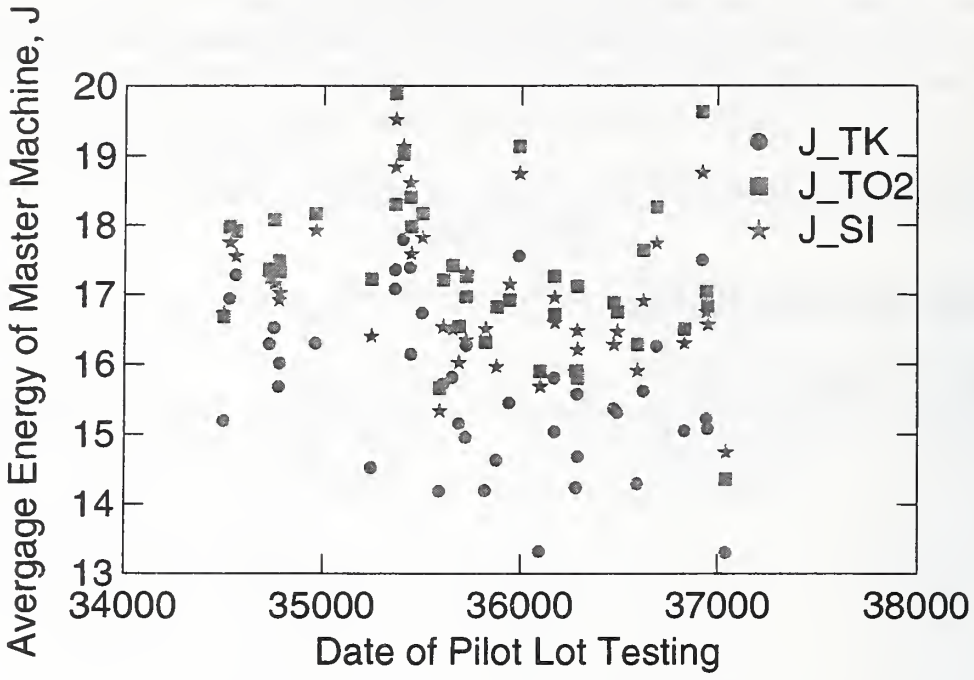


Figure A4.1.

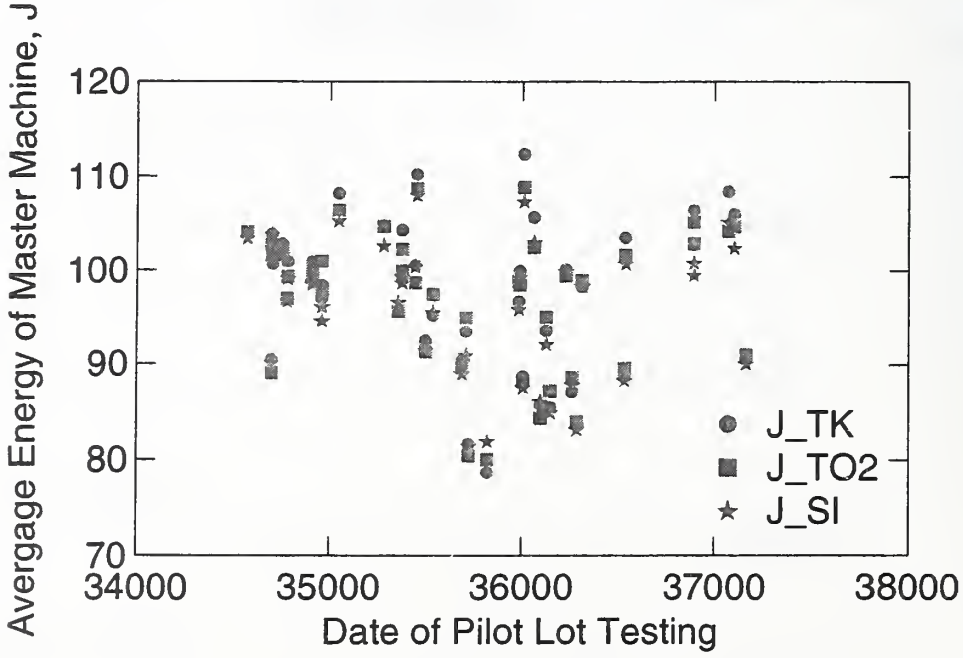


Figure A4.2.

Table A4.1.

LEVEL	Avg Energy, J	Energy_TK, J	Energy_TO, J	Energy_SI, J	STDPOOL, J	STDTK, J	STDTO, J	STDSI, J	CV
HIGH	103.89	104.08	104.10	103.47	2.34	2.65	2.48	3.21	0.02
HIGH	102.70	103.25	102.39	102.44	2.76	3.64	2.52	3.32	0.03
HIGH	89.87	90.48	89.03	90.11	2.14	2.04	2.73	2.08	0.02
HIGH	103.10	103.88	101.59	103.84	2.85	3.02	3.47	3.26	0.03
HIGH	101.09	100.71	101.67	100.90	3.31	3.85	3.87	3.13	0.03
HIGH	102.29	102.84	102.31	101.71	2.56	2.67	3.05	3.15	0.02
HIGH	98.20	100.94	96.99	96.74	2.60	3.53	2.38	2.21	0.03
HIGH	99.82	100.97	99.35	99.12	2.30	2.41	2.66	3.04	0.02
HIGH	99.76	100.90	99.57	98.81	2.73	3.56	2.43	3.74	0.03
HIGH	99.84	100.77	100.17	98.56	3.07	3.61	3.35	3.98	0.03
HIGH	97.55	97.03	101.00	94.63	2.76	3.67	2.60	2.53	0.03
HIGH	97.48	98.42	97.88	96.13	3.04	4.03	2.81	3.58	0.03
HIGH	106.67	108.23	106.44	105.35	3.04	3.96	2.85	3.92	0.03
HIGH	104.05	104.77	104.72	102.64	2.60	3.30	2.52	3.01	0.02
HIGH	95.99	95.67	95.67	96.63	2.45	2.54	3.02	2.48	0.03
HIGH	99.34	99.23	99.97	98.70	4.98	5.63	6.03	6.18	0.05
HIGH	102.00	104.32	102.30	99.38	6.02	6.93	7.27	7.86	0.06
HIGH	99.78	100.58	98.70	100.41	2.99	3.61	3.22	3.37	0.03
HIGH	109.01	110.23	108.78	108.03	3.41	3.46	4.41	3.46	0.03
HIGH	91.76	92.50	91.32	91.46	2.28	3.14	1.54	3.34	0.02
HIGH	96.07	95.21	97.47	95.53	2.78	3.26	3.00	3.53	0.03
HIGH	89.66	90.29	89.69	89.00	2.43	3.49	2.25	0.49	0.03
HIGH	93.14	93.51	94.96	90.96	3.19	3.86	3.50	3.30	0.03
HIGH	81.21	81.64	80.45	81.46	1.89	2.21	1.82	2.46	0.02
HIGH	80.05	78.69	80.02	81.95	3.40	3.83	4.02	3.72	0.04
HIGH	97.21	96.71	98.90	95.91	2.96	2.73	3.97	3.06	0.03
HIGH	99.26	99.98	98.46	99.34	1.89	2.19	1.86	2.43	0.02
HIGH	88.13	88.69	88.20	87.54	2.65	3.53	2.49	2.31	0.03
HIGH	109.52	112.37	108.87	107.38	3.25	4.27	3.12	3.80	0.03
HIGH	103.75	105.69	102.52	103.04	3.43	3.39	4.48	3.76	0.03
HIGH	85.34	85.62	84.32	86.09	2.88	2.98	3.63	2.88	0.03
HIGH	93.55	93.61	95.02	92.17	2.22	2.31	2.66	2.33	0.02
HIGH	85.82	85.46	87.17	84.94	2.08	1.72	2.78	2.31	0.02
HIGH	99.86	100.12	99.48	99.97	2.18	2.54	2.30	2.49	0.02
HIGH	87.86	87.10	88.61	87.84	2.52	3.29	2.45	2.30	0.03
HIGH	83.65	83.73	83.94	83.19	1.53	1.68	1.64	1.48	0.02
HIGH	98.57	98.34	99.00	98.37	2.17	2.36	2.39	2.87	0.02
HIGH	88.80	88.70	89.59	88.31	2.40	2.55	2.91	2.33	0.03
HIGH	101.96	103.51	101.71	100.75	2.94	2.95	3.62	4.22	0.03
HIGH	102.50	102.79	105.20	99.56	3.05	3.28	3.70	3.37	0.03
HIGH	106.97	106.42	102.93	100.84	3.37	3.69	4.18	3.09	0.03
HIGH	105.89	108.43	104.24	105.20	3.36	4.25	3.58	3.00	0.03
HIGH	104.40	106.03	104.75	102.47	2.71	2.96	3.29	2.46	0.03
HIGH	90.57	90.71	91.07	90.11	2.68	3.52	2.58	2.56	0.03
LOW	16.19	15.19	16.68	16.73	0.85	0.87	0.85	0.70	0.05
LOW	17.56	16.94	17.98	17.76	0.78	0.95	0.61	0.57	0.04
LOW	17.59	17.28	17.91	17.56	0.70	0.83	0.61	0.43	0.04
LOW	16.97	16.29	17.36	17.25	0.75	0.65	0.76	0.68	0.04
LOW	17.26	16.52	18.08	17.16	0.77	0.83	0.77	0.51	0.04
LOW	16.71	15.68	17.49	16.93	0.64	0.58	0.33	0.78	0.04
LOW	16.79	16.01	17.32	17.03	0.71	0.69	0.47	0.80	0.04
LOW	17.46	16.30	18.17	17.93	0.87	0.96	0.67	0.90	0.05
LOW	15.77	14.52	17.22	16.41	0.84	1.08	0.60	0.60	0.05
LOW	18.07	17.08	18.30	18.85	0.93	0.84	0.93	1.03	0.05
LOW	18.92	17.35	19.89	19.52	1.32	1.19	1.59	1.29	0.07
LOW	18.70	17.79	19.03	19.14	0.76	0.56	0.74	0.85	0.04
LOW	18.12	17.39	18.41	18.62	0.63	0.55	0.54	0.59	0.03
LOW	17.24	16.14	17.98	17.59	0.92	0.91	0.97	0.79	0.05
LOW	17.57	16.73	18.17	17.83	0.83	0.72	0.88	0.76	0.05
LOW	15.07	14.19	15.66	15.34	1.01	1.04	1.06	0.83	0.07
LOW	16.49	15.71	17.21	16.54	0.85	0.94	0.76	0.70	0.05
LOW	16.58	15.81	17.42	16.51	0.73	0.58	0.69	0.78	0.04
LOW	15.92	15.15	16.54	16.03	0.67	0.54	0.77	0.47	0.04
LOW	16.09	14.95	16.97	16.35	0.96	0.79	1.20	0.73	0.06
LOW	16.95	16.27	17.26	17.31	0.75	0.53	0.86	0.64	0.04

LOW	15.68	14.20	16.32	16.51	0.91	0.84	0.93	0.91	0.06
LOW	15.81	14.63	16.82	15.97	0.83	0.84	0.81	0.73	0.05
LOW	16.51	15.45	16.92	17.15	0.65	0.60	0.61	0.53	0.04
LOW	18.44	17.55	19.13	18.75	0.95	1.00	0.75	1.11	0.05
LOW	14.91	13.32	15.90	15.69	0.60	0.58	0.58	0.42	0.04
LOW	16.68	15.80	17.26	16.96	0.70	0.59	0.68	0.65	0.04
LOW	16.13	15.03	16.71	16.60	1.03	0.91	1.16	0.99	0.06
LOW	15.34	14.23	15.90	15.90	0.57	0.65	0.46	0.36	0.04
LOW	16.41	15.57	17.12	16.49	0.71	0.68	0.68	0.60	0.04
LOW	15.59	14.68	15.80	16.21	0.66	0.73	0.39	0.60	0.04
LOW	16.13	15.37	16.88	16.29	0.67	0.76	0.46	0.57	0.04
LOW	15.05	15.31	16.75	16.47	0.81	0.72	0.89	0.67	0.05
LOW	15.46	14.29	16.28	15.91	0.59	0.56	0.56	0.42	0.04
LOW	16.68	15.61	17.64	16.91	0.87	0.89	0.90	0.65	0.05
LOW	17.35	16.26	18.26	17.75	0.74	0.71	0.70	0.66	0.04
LOW	16.00	15.05	16.50	16.31	0.69	0.57	0.65	0.66	0.04
LOW	18.63	17.50	19.63	18.76	1.03	1.12	1.04	0.86	0.06
LOW	16.27	15.22	17.03	16.75	0.70	0.54	0.67	0.74	0.04
LOW	16.26	15.08	16.83	16.57	0.66	0.59	0.65	0.52	0.04
LOW	14.10	13.30	14.36	14.75	0.71	0.48	0.76	0.69	0.05
SUPERHI	222.52	223.38	219.04	225.14	6.98	8.69	7.93	7.76	0.03
SUPERHI	224.82	225.40	222.86	226.21	4.38	4.65	5.39	6.71	0.02
SUPERHI	230.78	228.58	229.89	233.86	4.02	5.22	3.85	6.48	0.02
SUPERHI	245.98	243.49	245.69	248.65	6.51	8.72	6.43	9.57	0.03
SUPERHI	227.09	224.65	227.88	228.59	6.45	9.35	5.46	7.71	0.03
SUPERHI	237.45	236.42	235.43	240.27	6.52	8.48	6.95	7.26	0.03
SUPERHI	245.45	245.30	244.25	246.91	5.98	8.08	5.82	8.21	0.02
SUPERHI	260.10	261.00	260.09	252.88	5.21	5.89	6.28	7.32	0.02
SUPERHI	209.92	203.05	219.72	206.87	6.89	8.86	7.58	6.42	0.03
SUPERHI	258.02	260.09	261.00	252.88	5.21	5.89	6.28	7.32	0.02
SUPERHI	209.88	203.05	219.72	206.87	6.89	8.86	7.58	6.42	0.03
SUPERHI	220.90	222.49	222.78	217.42	6.88	6.32	9.57	10.56	0.03
SUPERHI	227.29	226.12	225.68	230.07	7.27	8.64	8.71	8.10	0.03
SUPERHI	222.76	222.22	224.25	221.54	6.24	7.62	7.06	8.74	0.03
SUPERHI	222.76	222.22	224.25	221.54	6.24	7.62	7.06	8.74	0.03
SUPERHI	214.08	210.20	210.51	221.83	5.51	7.39	5.51	6.31	0.03
SUPERHI	205.95	206.54	200.70	210.49	5.35	6.61	6.10	5.14	0.03
SUPERHI	238.49	239.66	234.46	241.24	6.04	6.87	7.41	7.37	0.03
SUPERHI	200.79	200.19	199.41	202.69	6.23	8.37	6.26	7.21	0.03
SUPERHI	219.91	224.95	222.24	212.44	5.35	6.62	6.00	6.18	0.02
SUPERHI	220.86	221.38	220.89	220.15	6.52	8.70	6.78	5.86	0.03
SUPERHI	219.54	220.13	224.00	214.54	6.28	7.71	7.22	6.64	0.03

T. A. Siewert,[1] M. P. Manahan,[2] C. N. McCowan,[3] J. M. Holt,[4] F. J. Marsh,[5] and E. A. Ruth[6]

The History and Importance of Impact Testing*

Reference: Siewert, T. A., Manahan, M. P., McCowan, C. N., Holt, J. M., Marsh, F. J. and Ruth, E. A., **"The History and Importance of Impact Testing,"** *Pendulum Impact Testing: A Century of Progress, STP 1380*, T. A. Siewert and M. P. Manahan, Sr., Eds., American Society for Testing and Materials, West Conshohocken, PA, 2000, pp. 3-16.

Abstract: Charpy impact testing is a low-cost and reliable test method which is commonly required by the construction codes for fracture-critical structures such as bridges and pressure vessels. Yet, it took from about 1900 to 1960 for impact-test technology and procedures to reach levels of accuracy and reproducibility such that the procedures could be broadly applied as standard test methods. This paper recounts the early history of the impact test and reports some of the improvements in the procedures (standard specimen shape, introduction of a notch, correlation to structural performance in service, and introduction of shrouds) that led to this broad acceptance.

Keywords: absorbed energy; Charpy impact testing; impact testing; pendulum impact

Without uniformity of test results from day to day and from laboratory to laboratory, the impact test has little meaning. Over the years, researchers have learned that the results obtained from an impact test can depend strongly upon the specimen size and the geometry of the notch,

* Contribution of NIST; not subject to copyright.

[1] Supervisory Metallurgist, Materials Reliability Division, National Institute of Standards and Technology, Boulder, CO.

[2] President, MPM Technologies, Inc., 2161 Sandy Drive, State College, PA.

[3] Materials Research Engr., Materials Reliability Division, NIST, Boulder, CO.

[4] Alpha Consultants & Engineers, Pittsburgh, PA.

[5] Retired (Bethlehem Steel), San Marcos, CA.

[6] Tinius Olsen Test Machine Co., Willow Grove, PA.

Further details on the economic impact of Charpy impact testing are included in a previous version of this report published in *Standardization News*, February 1999.

anvils, and striker. To a lesser degree, impact test results also depend upon other variables such as impact velocity, energy lost to the test machine, and friction. The goal of those who have written and modified ASTM Standard Test Methods for Notched Bar Impact Testing of Metallic Materials (E 23) has over the years been to standardize and control the variables associated with impact testing. This report looks at the history of impact testing, with emphasis on the key advances in understanding and application of the impact test, as reflected in the evolution of the test standard.

Impact Testing: 1824 to 1895

The earliest publication we could find on the effects of impact loading on materials was a theoretical discussion by Tredgold in 1824 on the ability of cast iron to resist impulsive forces [*1*]. In 1849, the British formed a commission to study the use of iron in the railroad industry, which began by considering practical approaches to impact testing [*2*]. Apparently, failures of structures in the field were leading some researchers to speculate that impact loads affected materials far differently than static loads, so tensile-strength data (from slowly applied loads) was a poor predictor of performance under dynamic loads.

In 1857, Rodman devised a drop-weight machine for characterization of gun steels, and over the subsequent 30-year period, his machine was widely used to test railroad steels and for qualification of steel products [*2*]. Many of the early experiments with impact tests were performed on final product forms, such as pipes or axles. Thus they served as proof tests for a batch of material, or yielded comparative data for a new product design, or basic reference data on the impact resistance of different construction materials (such as the comparison of wrought iron to ductile iron). Instrumentation was poor for the early impact tests, so the data is often only as break or no-break for a mass dropped through a certain distance. These early drop weight tests were conducted using smooth (no notch or crack starter) rectangular bars. While the test worked well for brittle materials, where crack initiation is easy, specimens of ductile materials often just bent. LeChatalier introduced the use of notched specimens while conducting drop-weight tests in 1892 [*3*]. He found that some steels that showed ductile behavior (bending without fracture) in a smooth rectangular bar, would exhibit fragile behavior when the test specimen was notched. While the addition of a notch was a major improvement in the test method, a test procedure was needed that would provide a continuous, quantitative measure of the fracture resistance of materials. Also, substantial work was needed to develop test procedures that produced consistent data, and to answer the objections of those who doubted the value of impact testing.

1895 to 1922

This period saw the establishment of a number of national and international standards bodies, which took up the causes of developing robust test procedures and developing consensus standards for many technologies, including impact testing. One of these standards bodies was The American Society for Testing and Materials, established in 1898. Another was the International Association for Testing Materials, officially established in 1901, but this association grew out of the good response to two previous International Congresses that had been held a number of years before. These two standards bodies seem to have had a good working

relationship, and the President of ASTM, Prof. H. M. Howe, also served on the Board of IATM during this time [*4*].

In 1902, only four years after the founding of ASTM, the ASTM "Committee on the Present State of Knowledge Concerning Impact Tests" published a bibliography on impact tests and impact testing machines in the second volume of the Proceedings of ASTM [*5*]. This bibliography listed more than 100 contemporary papers on impact testing published in the U.S., France, and Germany. Many of these papers contained information that was also known to the members of IATM. In fact, some of the papers had been presented and discussed at the IATM Congresses.

Among the references is a report by Russell (published in 1898) that shows remarkable insight into the needs of the design engineers of the time and introduces quantitative measurement to the test [*6*]. He pointed out that none of the machines of the time, typically of a drop-weight design, had the ability to determine any data beyond whether the specimen broke or remained intact. Therefore, he designed and built a pendulum machine which "would measure the energy actually absorbed in breaking the test bar". His report shows a test machine that is based on the same swinging pendulum concept as those in common use today and mentions his careful analysis of the mechanics of the test, including corrections for friction losses and calculation and comparison of the centers of gravity and percussion. Since this was before the time of compact, standardized test specimens, the machine was vary large and massive, and was capable of breaking many full-size products. Besides showing a prototype of the machines used today, this report is valuable in that it includes data on over 700 tests of typical construction materials, and emphasizes the effect of the rate of loading in evaluating materials for different service conditions. Russell's pendulum impact machine finally provided a means for quantifying the energy absorbed in fracturing a test specimen for a wide range of materials and conditions. His paper nicely summarizes the test-machine technology and knowledge for material performance at the end of the past century, and so served as a benchmark for future research. To the best of our knowledge, Russell was the first to develop and demonstrate the advantages of the pendulum design for impact testing machines.

The members of IATM Commission 22 (On Uniform Methods of Testing Materials) continued to conduct research that addressed the shortcomings in the impact testing techniques, until they had developed a knowledge of most of the important factors in the test procedure. Even though many of these early machines and reports are simplistic by today's standards, they provided previously unknown data on the impact behavior of materials. France seems to have been an early adopter of impact testing for infrastructure construction standards, and so French researchers provided much data on the effects of procedure variables and were the most prolific contributors to the IATM Proceedings between 1901 and 1912. Incidentally, it was a representative from France, G. Charpy, who became the chair of the impact testing activity after the 1906 IATM Congress in Brussels, and presided over some very lively discussions on whether impact testing procedures would ever be sufficiently reproducible to serve as a standard test method [*7*]. Charpy's name seems to have become associated with the test because of his dynamic efforts to improve and standardize it, both through his role as Chairman of the IATM Commission and through his personal research [*8*]. He seems to have had a real skill for recognizing and combining key advances (both his and those of other researchers) into

continually better machine designs and consensus procedures. For example, Charpy acknowledges the benefits of Russell's pendulum design in his 1901 paper [*8*] by stating: "Russell described in a paper presented in 1897 at the American Society of Civil Engineers some 'experiments with a new machine for testing materials by impact.' The machine he is using is designed to determine the work absorbed by the rupture of a bar, for this, the ram used appears in the form of a pendulum arranged in such a way so that when it is released from its equilibrium position, it meets the test bar in passing through the vertical position, breaks it and afterward rises freely under the influence of the acquired speed. The difference between the starting height and the finishing height of the pendulum allows evaluation of the work absorbed by the rupture of the bar."

By 1905, Charpy had proposed a machine design that is remarkably similar to present designs and the literature contains the first references to "the Charpy test" and "the Charpy method". He continued to guide this work until at least 1914 [*7,9-10*]. A number of other standard machine designs and procedures were also under consideration at this time, and in 1907 the German Association for Testing Materials adopted one developed by Ehrensberger [*10*]. Because the pendulum machine had not achieved dominance yet, impact machine designers and manufacturers offered three major types; Drop Weight (Fremont, Hatt-Turner, and Olsen), Pendulum Impact (Amsler, Charpy, Dow, Izod, Olsen, and Russell), and Flywheel (Guillery).

This was a period during which the configuration and size of specimens closely approached what we use today [*7*]. Originally, two standard specimen sizes were most popular. The smaller had a cross section of 10 by 10 mm, a length of about 53 mm (for a distance of 40 mm between the points of support), a notch 2 to 5 mm deep, and a notch tip radius near 1 mm. The larger and initially more popular of these specimen sizes was scaled up by a factor of three in all these dimensions. The group favoring the larger specimen pointed out the advantage of sampling a larger cross section of the material (for reduced scatter in the data) and the difficulty of producing the small notch radius on the smaller specimen. However, the group favoring the smaller specimen eventually won because a more compact and lower-cost machine could be used, and not all structures were thick enough to produce the larger specimen. Besides specimen dimensions that are very similar to what we use today, the Commission proposed features for a standard impact procedure that included:

- limits for the velocity of the striker,
- rigid mounting to minimize vibration losses,
- a minimum ratio of anvil mass and rigidity to striker size, and
- recognition of the artificial increase in energy as ductile specimens deform around the edges of a wide striker [*7*].

One report at the 1912 meeting [*7*] included the testimonial from a steel producer of how the improved impact test procedures had allowed them to tailor the refining processes to produce less brittle steel. The report describes a reduction by a factor of 20 in the number of production parts that were rejected for brittle performance.

1922 to 1933: The Beginning of ASTM Method E 23

ASTM Committee E-1 on Methods for Testing sponsored a Symposium in 1922 on Impact Testing of Materials as a part of the 25th Annual Meeting of the Society, in Atlantic City, New Jersey. The Symposium included a history of the developments in this area, a review of work done by the British Engineering Standards Association, several technical presentations, and the results of a survey sent to 64 U.S. testing laboratories [*11*]. Twenty-three respondents to the survey offered detailed information on topics such as the types of machines in use, the specimen dimensions, and procedures. In addition, many responded positively to a question about their willingness to develop an ASTM standard for impact testing.

Based on the information in this survey, an ASTM subcommittee began to prepare a standard test method for pendulum impact testing in 1923. This effort took until 1933, when ASTM published "Tentative Methods of Impact Testing of Metallic Materials," ASTM designation E 23-33T. (An ASTM specification of "Tentative" indicated that it was subject to annual review and was a work in progress. The tentative designation is no longer used by ASTM.) (Other countries also developed their own standards; however, we found it difficult to find their records and to track their developments.)

ASTM E 23-33T specified that a pendulum-type machine was to be used in testing and "recognized two methods of holding and striking the specimen", that is, the Charpy test and the Izod test (where the specimen is held vertically by a clamp at one end). It did not specify the geometry of the striking edge (also known at the time as the "tup") for either test. It stated that "the Charpy type test may be made on unnotched specimens if indicated by the characteristics of the material being tested, but the Izod type test is not suitable for other than notched specimens". Only a V-notch was shown for the Charpy test. Although the dimensions for both types of specimens were identical with those currently specified, many tolerances were more restrictive. The units were shown as English preferred, metric optional. The committee pointed out many details that influence the test results, but because they did not have the knowledge and database needed to specify values and/or tolerances for these details, the document was issued as a tentative. The original document contains an appendix with general discussions of applications, the relation to service conditions, and comparisons between materials. As our understanding of the variables in Charpy testing has grown, ASTM E 23 has been revised repeatedly to incorporate the new knowledge.

1934 to 1940

The first revision of E 23 was issued in 1934 and it added a dimension for the radii of the anvil and specifically stated that "these specimens (both the Charpy and the Izod) are not considered suitable for tests of cast iron" referencing a report of ASTM Committee A3 on Cast Iron. The method retained the "tentative" designation.

The geometry of the Charpy striking tup, specifically the radius of the tup that contacted the specimen, was not specified in the 1934 revision. However, the minutes of the 1939 and 1940 meetings for the Impact Subcommittee of E1 state that this item was discussed and a survey was

made of the geometries used in the United Kingdom and in France. Those countries had been using radii of 0.57 mm and 2 mm, respectively. For reasons that were not recorded, the members of the Subcommittee agreed to a radius of 8 mm at the 1940 meeting and ASTM E 23 was revised and reissued as E 23-41T. Two other changes that occurred with this revision were that metric units became the preferred units, and keyhole and U notches were added for Charpy-test specimens.

1940 to 1948

Impact testing seems to have been a useful technique for evaluating materials, but was not a common requirement in purchase specifications and construction standards until the recognition of its ability to detect the ductile-to-brittle transition in steel. Probably the greatest single impetus toward implementation of impact testing in fabrication standards and material specifications came as a result of the large number of ship failures that occurred during World War II. These problems were so severe that the Secretary of the U.S. Navy convened a Board of Investigation to determine the causes and to make recommendations to correct them. The final report of this Board stated that of 4694 welded-steel merchant ships studied from February 1942 to March 1946, 970 (over 20%) suffered some fractures that required repairs [*12*]. The magnitudes of the fractures ranged from minor fractures that could be repaired during the next stop in port, to 8 fractures that were sufficiently severe to force abandonment of these ships at sea. Remedies included changes to the design, changes in the fabrication procedures and retrofits, as well as impact requirements on the materials of construction. The time pressures of the war effort did not permit thorough documentation of the effect of these remedies in technical reports at that time; however, assurance that these remedies were successful is documented by the record of ship fractures that showed a consistent reduction in fracture events from over 130 per month in March 1944 to less than five per month in March 1946, even though the total number of these ships in the fleet increased from 2600 to 4400 during this same period [*12*].

After the war, the National Bureau of Standards released its report on an investigation of fractured plates removed from some of the ships that exhibited these structural failures and so provided the documentation of the importance of impact testing [*13*]. The NBS study included chemical analysis, tensile tests, microscopic examination, Charpy impact tests, and reduction in thickness at the actual ship fracture plane. A notable conclusion of the report was that the plates in which the fracture arrested had consistently higher impact energies and lower transition temperatures than those in which the fractures originated. This was particularly important because there was no similar correlation with chemical composition, static tensile properties (all steels met the ABS strength requirements), or microstructure. In addition, the report established 15 ft-lb (often rounded to 20 J for metric requirements) as a minimum toughness requirement, and recommended that "some criterion of notch sensitivity should be included in the specification requirements for the procurement of steels for use where structural notches, restraint, low temperatures, or shock loading might be involved," leading to a much wider inclusion of Charpy requirements in structural standards.

1948 to Present

By 1948, many users thought that the scatter in the test results between individual machines could be reduced further, so additional work was started to more carefully specify the test method and the primary test parameters. By 1964, when the ASTM E 23 standard was revised to require indirect verification testing, the primary variables responsible for scatter in the test were well known. In a 1961 paper, Fahey [*14*] summarized the most significant causes of erroneous impact values as follows: (1) improper installation of the machine, (2) incorrect dimensions of the anvil supports and striking edge, (3) excessive friction in moving parts, (4) looseness of mating parts, (5) insufficient clearance between the ends of the test specimen and the side supports, (6) poorly machined test specimens, and (7) improper cooling and testing techniques. While the machine tolerances and test techniques in ASTM E 23 addressed these variables, it was becoming apparent that the only sure method of determining the performance of a Charpy impact machine was to test it with standardized specimens (verification specimens).

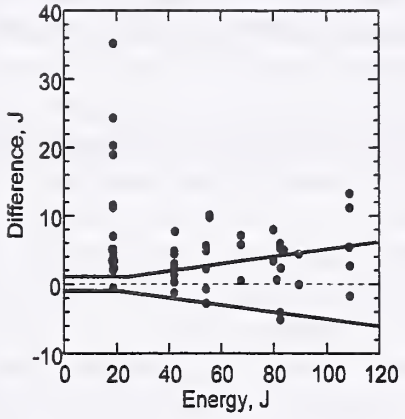


Figure 1. The deviation and energy values obtained for the first round of tests on industrial machines. The deviation is calculated as the difference between the results of the Watertown Arsenal machines and the industrial machines. These data were originally published by D.E. Driscoll, Reproducibility of Charpy Impact Test, ASTM STP 176, 1955.

Much of the work that showed impact tests did not have inherently high scatter, and could be used for acceptance testing, was done by Driscoll at the Watertown Arsenal [*15*]. Driscoll's study set the limits of 1 ft-lb (1.4 J) and ± 5%, shown in **Figures 1 and 2**. The data superimposed on these limits in **Figures 1 and 2** are the initial verification results gathered by Driscoll for industrial impact machines to evaluate his choice of verification limits.

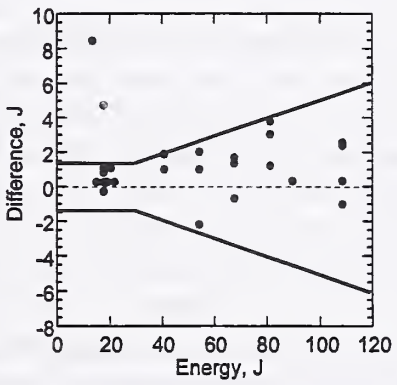


Figure 2. The deviation and energy values for the second and third rounds of tests on industrial machines. The data show that all but two of the machines tested were able to pass the 1.4 J or 5% criteria after appropriate repairs were made. These data were originally published by D. E. Driscoll, Reproducibility of Charpy Impact Test, ASTM STP 176, 1955.

In **Figure 1**, the verification results for the first attempt on each machine are shown: only one machine fell within the ± 1 ft-lb (1.4 J) limit proposed for the lower energy range. Results for retests on the same machines after maintenance are shown in **Figure 2**. Driscoll's work showed the materials testing community that not all machines in service could perform well enough to meet the indirect verification requirements, but that most impact machines could meet the proposed requirements if the test was conducted carefully and the machine was in good working

condition. With the adoption of verification testing, it could no longer be convincingly argued that the impact test had too much inherent scatter to be used as an acceptance test.

Early results of verification testing showed that 44% of the machines tested for the first time failed to meet the prescribed limits, and it was thought that as many as 50% of all the machines in use might fail [*16*]. However, the early testing also showed that the failure rate for impact machines would drop quickly as *good* machines were repaired, *bad* machines were retired, and more attention was paid to testing procedures. It was estimated that approximately 90% of the machines in use could meet the prescribed limits of ± 1 ft-lb (1.4 J) or ± 5%. Recently acquired verification specimen data, shown in **Figures 3 through 5**, confirm these predictions. Failure rates for verification tests at low, high, and super-high energy ranges are currently estimated to be 12, 7, and 10%, respectively [*17*].

Overall, the incorporation of verification limits in ASTM E 23 has greatly improved the performance of impact machines, so that data collected using ASTM E 23 machines can be compared with confidence. ASTM E 23 is still the only standard in the world, to our knowledge, that requires very-low-energy impact specimens (between 15 and 20 J) for verification, and as shown by the data in **Figure 1**, results obtained using machines in need of maintenance can vary by more than 100% at this energy level. In effect, the limits imposed by ASTM E 23 have produced a population of impact machines that are arguably the best impact machines for acceptance testing in the world.

While ASTM E 23 is used around the world, there are other forums for the development of global standards. One of these, the International Organization for Standardization, ISO, allows qualified representatives from all over the world to come together as equal partners in the resolution of global standardization problems [*18*]. ISO Committee TC 164 handles the topic of Mechanical Testing, and its Subcommittee SC 4 handles toughness testing. While this

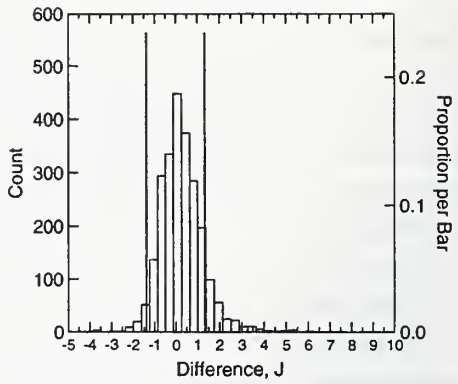


Figure 3. Distribution of low-energy verification data. Data for 1995-1997. Approximately 2400 tests; each test is an average for five specimens. The vertical lines at ±1.4 J represent the acceptance criteria.

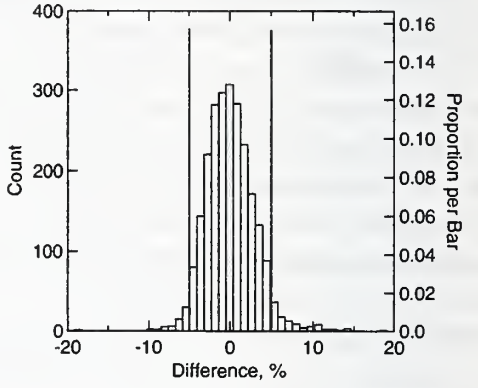


Figure 4. Distribution of high-energy verification data. Data for 1995-1997. Approximately 2400 tests. Each test is an average for five specimens. The vertical lines at ±5% represent the acceptance criteria.

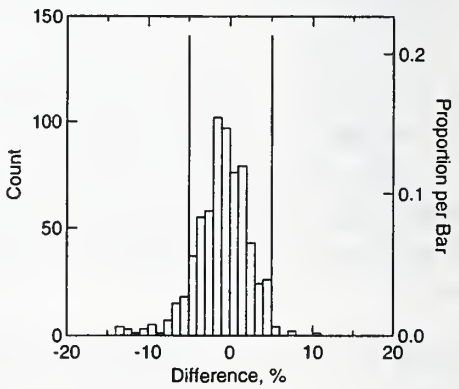


Figure 5. Distribution of the super-high-energy verification data. Data for 1995-1997. Approximately 650 tests. Each test is an average for five specimens. The vertical lines at ±5% represent the acceptance criteria.

subcommittee has developed and maintains ten standards on toughness testing, perhaps the most pertinent is ISO Standard R 442:1965 *Metallic Materials - Impact Testing - Verification of Pendulum Impact Machines*. This standard covers the Charpy test and is presently undergoing balloting for revision. An important feature of this document is that it recognizes Charpy testing with both the 2-mm and 8-mm radius striker. There are other regional and national standards that specify impact testing procedures, such as the Japanese standard, JIS Z2242, *Method for Impact Test for Metallic Materials*.

Typical Applications Today

Nuclear

Since it is impractical to measure the fracture toughness of large specimens throughout the life of a nuclear power plant, surveillance programs use Charpy and tensile specimens to track the embrittlement induced by neutrons. The economic importance of the Charpy impact test in the nuclear industry can be estimated by noting that most utilities assess the outage cost and loss of revenue for a nuclear plant to be in the range of $300,000 to $500,000 per day. If Charpy data can be used to extend the life of a plant one year beyond the initial design life, a plant owner could realize revenues as large as $150,000,000. Further, the cost avoidance from a vessel related fracture is expected to be in the billion-dollar range. To date, the NRC has shut down one U.S. plant as a result of Charpy data trends. It is important to note that this plant's pressure vessel was constructed from a one-of-a-kind steel and is not representative of the U.S. reactor fleet.

Nonetheless, with decisions like this based on the Charpy test, the importance of ASTM E 23 and the restraints it applies cannot be overemphasized.

Steel

The Charpy V-notch (CVN) test specimen and associated test procedure is an effective cost-saving tool for the steel industry. The specimen is relatively easy to prepare, many specimens can be prepared at one time, various specimen orientations can be tested, and relatively low-cost equipment is used to test the specimen. In many structural steel applications, the CVN test can be used: (1) as a quality control tool to compare different heats of the same type of steel, (2) to check conformance with impact requirements in standards, and (3) to predict service performance of components. Also, CVN test information can be correlated with fracture toughness data for a class of steels so that the results of fracture-mechanics analyses can be compared with the material toughness.

CVN data have many uses, such as during the design and construction of a bridge or an offshore oil platform. Before full-scale production of the steel order can begin, the supplier needs to demonstrate to the buyer that the steel plate is capable of meeting certain design criteria. The process begins by making the steel grade and then testing a portion of the plate to determine if all required criteria are met. Also, steel mill equipment imposes limitations on plate size; therefore, individual steel plates need to be welded together in the field to produce lengths which

can reach deep into ocean waters. Small sections of the sample plate are welded together, and fracture mechanics tests are conducted to determine the crack tip opening displacement (CTOD) toughness in the heat affected zone (HAZ) and in areas along the fusion line where the weld metal meets the base metal. Then, a steel supplier might correlate the CTOD test results with CVN 50% ductile-brittle transition temperature (DBTT). By agreement between the customer and supplier, this correlation can allow the steel supplier to use the Charpy test instead of the more expensive and time-consuming CTOD testing.

Continuing Standardization Efforts

Even after 100 years, the Charpy impact test procedures still have room for improvement. The ASTM E 23 standard has recently been redrafted to provide better organization and to include new methods such as in-situ heating and cooling of the test specimens. Two new related standards are also under development through ASTM Task Group E 28.07.08, "Miniature and Instrumented Notched Bar Testing", which was formed a little more than two years ago. The first standard covers miniature notched bar impact testing and relies on many of the existing practices related to test machine requirements and verification as specified in existing standard E 23. The second standard is focused on instrumented testing, where strain gages attached to the striker provide a force-deflection curve of the fracture process for each specimen. Research is focused on using these data to obtain plane strain fracture toughness as well as other key test parameters. Upon acceptance of the standard by ASTM, both the existing E 23 standard and the new miniature notched bar standards would reference the instrumented impact standard.

The state of the art in impact testing continues to advance in other parts of the world also. ISO is balloting a standard (14556) on instrumented impact testing, there is work in Europe on miniature Charpy specimens, and ESIS is investigating the use of pre-cracked Charpy specimens for determining fracture toughness. It can be expected that harmonization efforts will bring some of this work into E 23 in the future.

Conclusion

Several years ago, at the ASTM Symposium on "The Charpy Impact Test: Factors and Variables" [*19*], a bystander was overheard to say: "I see that there is a Symposium on the Charpy Test; what can be new there?" Since then, the document has been updated twice and is currently being revised to reflect new developments and to make it more "user friendly." Although ASTM E 23 has been a useful standard for many years, it continues to be a "work in progress," a work used extensively to help evaluate existing and new materials for products and structures -- a test to ensure safety as well as to reduce the initial and lifetime costs for structures. The ASTM E 23 standard is a document that continues to improve as our technical knowledge increases. Knowledge which will help make the test more accurate and reliable is continually being gained. New technologies such as miniaturization of the test, instrumenting the striker to obtain additional data, and developing mechanics models to enable extraction of plane strain fracture toughness will be areas of development over the next 100 years. We anticipate that the benefits from the application of E 23 during the next 100 years will overshadow the benefits from those in the past 100 years.

References

[1] Tredgold, T., *Strength of Cast Iron*, 1824, pp. 245-268.

[2] White, A.E. and Clark, C.L., *Bibliography of Impact Testing*, Department of Engineering Research, University of Michigan, 1925.

[3] LeChatalier, A., "On the Fragility After Immersion in a Cold Fluid", *French Testing Commission*, Volume 3, 1892.

[4] "Report on the Work of the Council from the Budapest to the Brussels Congress - 1901-1906," *Proceedings of the International Association for Testing Materials, Brussels Congress,* 1906.

[5] Hatt, W.K. and Marburg, E., "Bibliography on Impact Tests and Impact Testing Machines," *Proceedings ASTM*, Vol. 2, 1902, p. 283.

[6] Russell, S. B. "Experiments with a New Machine for Testing Materials by Impact," *Transactions ASCE*, Vol. 39, June 1898, p. 237.

[7] *Proceedings of the Sixth Congress of the International Association for Testing Materials*, New York, 1912.

[8] Charpy, M.G., "Note sur l'Essai des Metaux a la Flexion par Choc de Barreaux Entailles, *Soc. Ing. Francais*, June 1901, p. 848.

[9] "Impact Testing of Notched Bars," *The Engineer*, Vol. 99, March 10, 1905 pp. 249-250.

[10] Whittemore, H. L., "Resume of Impact Testing of Materials, with Bibliography," *Proceedings ASTM*, Vol. 22, Part 2, 1922, p. 7.

[11] Warwick, C. L., "Resume on Notched Bar Tests of Metals," *Proceedings of ASTM*, Vol. 22, Part 2, 1922, p. 78.

[12] The Design and Methods of Construction of Welded Steel Merchant Vessels: Final Report of a (U.S. Navy) Board of Investigation, *Welding Journal*, Vol. 26, No. 7, July 1947, p. 569.

[13] Williams, M. L. and Ellinger, G. A., *Investigation of Fractured Steel Plates Removed from Welded Ships*, National Bureau of Standards Report, December 9, 1948.

[14] Fahey, N. H., "Effects of Variables in Charpy Impact Testing," *Materials Research Standards*, Vol. 1, No. 11, Nov., 1961.

[15] Driscoll, D. E., "Reproducibility of Charpy Impact Test," *Impact Testing, ASTM STP 176*, 1955.

[16] Fahey, N. H., "The Charpy Impact Test - Its Accuracy and Factors Affecting Test Results," *Impact Testing of Metals, ASTM STP 466*, ASTM, 1970.

[17] McCowan, C. N., Wang, C. M., and Vigliotti, D. P., "Summary of Charpy Impact Verification Data: 1994 - 1996," Submitted to the *Journal of Testing and Evaluation*, 1998.

[18] More information is on the ISO World Wide Web site, at http://www.iso.ch

[19] *Charpy Impact Test: Factors and Variables, ASTM STP 1072*, J. M. Holt, Ed., ASTM, 1990.

M.P. Manahan, C.N. McCowan, T.A. Siewert, J.M. Holt, F.J. Marsh, and E.A. Ruth

For 100 Years, Notched Bar Impact Testing Standards Have Yielded Widespread Benefits for Industry

ASTM Standardization News, 27(2), February 1999, pp.30-35.

Without standardization, the impact test has little meaning. The test result obtained from an impact test is dependent upon the specimen size, notch geometry, and the geometry of the anvils and striker. To a lesser degree, impact test results are also dependent upon other variables such as impact velocity, energy lost to the test machine, and friction. The goal of those who have written and modified ASTM E 23, Standard, Test Methods for Notched Bar Impact Testing of Metallic Materials, has been, and is, to standardize and control the variables associated with impact testing.

ASTM E 23 describes notched-bar impact testing of metallic materials by both the Charpy (simple-beam) and the Izod (cantilever-beam) methods. While for some materials the Izod method is used, the Charpy method is the overwhelming choice for metallic materials. Because it is the method of choice, ASTM E 23 and this paper give more attention to the Charpy method.

In the late 1800s and early 1900s, many different specimen and anvil geometries were being used for Charpy testing. As a result it was difficult, if not impossible, to correlate the results of various investigators, and the energy absorbed in fracturing a Charpy test specimen was in question. In 1909, Révillon concluded that standardizing the specimen and anvil geometries was necessary to obtain reproducible results and thus overcome the opposition to impact tests.[1] The first version of ASTM E 23 accomplished that. Various manufacturers of impact machines could now make machines that would produce test results similar to tests performed on their competitors' machines. Further revisions to ASTM E 23 strived to make the impact test result a more accurate measure of the energy required to fracture the specimen. These changes included: (1) modifications to the striker, (2) standardizing windage and friction losses and corrections, (3) standardizing requirements for test machine foundations, (4) standardizing specimen supports and shrouds to reduce "jamming" of the broken specimen, which falsely increases the measured energy, (5) standardizing the impact velocity, and (6) requiring verification of sources of test errors.

As these changes were made to ASTM E 23, manufacturers of impact test machines and users of the test methods were forced to change their designs and procedures in order to conform to the standard. As a result, impact tests performed today, in accordance with ASTM E 23, are reliable and reproducible throughout the world.

The Early Days of Impact Testing

The earliest known publication on the effects of impact loading on materials was a theoretical discussion by Tredgold in the early 1800s on the ability of cast iron to resist impulsive forces.[2] In 1849, the British formed a commission to study the use of iron in the railroad industry, which began its work by considering practical approaches to impact testing.[3] In 1857, Captain Rodman [2] devised a drop weight machine for characterization of gun steels, and over the subsequent 30-year period, this machine was widely used in the testing of railroad steels and for qualification of steel products. These early drop-weight tests were conducted using smooth (no notch or crack starter) rectangular bars. While the test worked well for brittle materials, where initiating a crack is easy, ductile materials would only bend and inducing fracture was not possible. LeChatalier introduced the use of notched specimens while conducting drop weight tests in 1892.[4] It was discovered that some steels that showed ductile behavior (bending without fracture) in a smooth rectangular bar, would exhibit fragile behavior when the test specimen was notched. While the addition of a notch was a major improvement in the test method, a test procedure was needed that would provide a continuous, quantitative measure of the fracture resistance of materials.

Pendulum Impact Testing

A hundred years ago, in 1898, a report by Russell showed a machine that is based on the same swinging-pendulum concept as those in common use today.[5] His report included data on many construction materials, and emphasized the effect of the rate of loading in evaluating materials for different service conditions. The pendulum impact machine of Russell finally provided a means for measuring the energy absorbed in fracturing a test specimen for a wide range of materials and conditions, from brittle at low test temperatures to ductile at high test temperatures.

Impact testing was an exciting and active research field near the turn of the century and a 1902 bibliography listed more than 100 contemporary papers on impact testing published in the U.S., France, and Germany.[6] Such studies were compared and discussed at the meetings of the International Association for Testing Materials (before ASTM took up this topic) during the next decade.[7,8] During this period, the committee members conducted research that overcame the shortcomings in the impact testing techniques, until they had developed robust and carefully considered procedures that provided useful information for industrial users. Even though these early standardized procedures were primitive by today's standards, they proved very satisfactory in evaluating the impact behavior of materials. For example, these early reports record that the test procedures were adopted by the French Navy for ship machinery, especially for engine shafting. Incidentally, it was a representative from France, G. Charpy, who became the chair of the impact testing activity after the 1906 IATM Congress in Brussels, and presided over some very lively discussions on whether impact testing procedures would ever be sufficiently reproducible to serve as a standard test method.[9] Although not the inventor of the pendulum impact test, Charpy's name is associated with the test because of his efforts to improve and standardize it.

Development of ASTM Method E23

In about 1923, an ASTM subcommittee began to prepare a standard test method for pendulum impact testing. This effort took until 1933 when ASTM published "Tentative Methods of Impact Testing of Metallic Materials, " ASTM designation E 23-33T. (An ASTM specification of "Tentative" indicated that it was subject to annual review and was a work in progress. The tentative designation is no longer used by ASTM.)

Method E23-33T specified that a pendulum-type machine was to be used in testing and "recognized two methods of holding and striking the specimen," that is, the Charpy test and the Izod test. It did not specify the geometry of the striking edge (also known as the "tup") for either test. It stated that "the Charpy type test may be made on unnotched specimens if indicated by the characteristics of the material being tested, but the Izod type test is not suitable for other than notched specimens." Only a V-notch was shown for the Charpy test. Although the dimensions for both types of specimens were identical with those currently specified, many tolerances were more restrictive. The units were shown as English preferred, metric optional. The authors pointed out many details that influence the test results, but because they did not have the knowledge to specify values and/or tolerances for these details, the document was issued as a tentative one. The original document contains an appendix with general discussions of applications, the relation to service conditions, and comparisons between materials. As our understanding of the Charpy-test variables has grown, Method E23 has been revised to incorporate the new knowledge.

The first revision was issued in 1934 and added a dimension for the radii of the anvil and specifically stated that "these specimens (both the Charpy and the Izod) are not considered suitable for tests of cast iron" referencing a report of ASTM Committee A3 on Cast Iron. The method retained the "tentative" designation.

The geometry of the Charpy striking tup, specifically the radius of the tup that contacted the specimen, was not specified in the 1934 revision, but the minutes of the 1939 and 1940 meetings for the Impact Subcommittee E1 state that this item was discussed and a survey was made of the geometries used in the United Kingdom and in France. Those countries used radii of 0.57 mm and 2 mm, respectively. For reasons that were not recorded, the members of the Subcommittee agreed to a radius of 8 mm at the 1940 meeting and ASTM E23 was revised and reissued as E 23-41T. Two other changes that occurred with this revision were that metric units became the preferred units and keyhole and U notches were added for Charpy-test specimens.

Impact testing seems to have been a useful material evaluation technique, but was not a common requirement in purchase specifications and construction standards until the recognition of its ability to detect the ductile-to-brittle transition in steel. Probably the greatest single impetus toward implementation of impact testing in fabrication standards and material specifications came as a result of the large number of ship failures that occurred during World War II. These problems were so severe that the Secretary of the Navy convened a Board of Investigation to determine the causes and to make recommendations to correct them. The final report of this Board summarized the magnitude of the problems found during this study.[10] Of

4,694 welded-steel merchant ships studied from February 1942 to March 1946, 970 (over 20 percent) suffered some fractures that required repairs. The magnitudes of the fractures ranged from minor fractures that could be repaired during the next stop in port, to eight fractures that were sufficiently severe to force abandonment of these ships at sea. Also, at least 26 lives were lost because of these fractures. The total cost of these fractures to our nation (replacement and repair of ships, loss or delay of critical cargo needed for the war effort, and loss of lives) was immense.

The problem was complex and remedies included changes to the design, changes in the fabrication procedures, retrofits, as well as impact requirements on the materials of construction. Assurance that these remedies were successful is documented by the record of ship fractures that showed a consistent reduction in fracture events from over 130 per month in March 1944 to less than 5 per month in March 1946, even though the total number of these ships in the fleet increased from 2,600 to 4,400 during this same period.[10]

Benefits from the Introduction of Verification Testing

In 1948, many users felt that the scatter in the test results between individual machines could be reduced further, so additional work was started to more carefully specify the test method and the primary test parameters. By 1964, when the ASTM E 23 standard was revised to require indirect verification testing, the primary variables responsible for scatter in the test were well known. In a 1961 paper, Fahey[11] summarized the most significant causes of erroneous impact values as follows: (1) improper installation of the machine, (2) incorrect dimensions of the anvil supports and striking edge, (3) excessive friction in moving parts, (4) looseness of mating parts, (5) insufficient clearance between the ends of the test specimen and the side supports, (6) poorly machined test specimens, and (7) improper cooling and testing techniques. While the machine tolerances and test techniques in ASTM E 23 addressed these variables, it was becoming apparent that the only sure method of determining the performance of a Charpy impact machine was to test it with standardized specimens (verification specimens).

Much of the work that showed impact tests did not have inherently high scatter, and could be used for acceptance testing, was done by Driscoll at the Watertown Arsenal.[12] Driscoll's study set the limits 1 ft-lb and ±5 percent, shown in Figures 1 and 2. The data superimposed on these limits in Figures 1 and 2 are the initial verification results gathered by Driscoll for industrial impact machines to evaluate his choice of verification limits. In Figure 1, the verification results for the first attempt on each machine are shown: only one machine fell within the ± 1 ft-lb limit proposed for the lower energy range. In Figure 2, results for re-tests on the same machines are shown, after maintenance. Driscoll's work showed the materials impact testing community that not all machines in service could perform well enough to meet the indirect verification requirements, but that most impact machines could meet the proposed requirements if the test were conducted carefully and the machine were in good working condition. With the adoption of verification testing, it could no longer by convincingly argued that the impact test had too much inherent scatter to be used as an acceptance test.

Early verification test results showed that 44 percent of the machines tested for the first time failed to meet the prescribed limits, and it was thought that as many as 50 percent of all the machines in use might fail.[13] However, the early testing also showed that the failure rate for impact machines would drop quickly as good machines were repaired, bad machines were retired, and more attention was paid to testing procedures. It was estimated that approximately 90 percent of the machines in use could meet the prescribed limits of ± 1 ft-lb or ± 5 percent. Recently acquired verification specimen data, shown in **Figures 3 through 5**, confirms these predictions. Failure rates for verification tests at low, high, and super-high energy ranges are currently estimated to be 12, 7, and 10 percent respectively.[14]

Overall, the impact of incorporating verification limits in ASTM E 23 has greatly improved the performance of impact machines; data collected using ASTM E 23 machines can be compared with confidence. ASTM E 23 is still the only standard in the world, to our knowledge, that requires low-energy impact tests for verification, and as shown by the data in **Figure 1**, results obtained using machines in need of maintenance can vary by more than 100 percent at the low-energy level. In effect, the limits imposed by ASTM E 23 have produced a population of impact machines that are arguably the best impact machines for acceptance testing in the world.

Case Study In the Nuclear Industry

Notched-bar impact data do not directly provide engineering data (such as plane strain fracture-toughness data) which can be used in structural integrity analyses. Therefore, it is necessary to conservatively correlate the key test parameters with component performance objectives. An important illustration of this approach is the nuclear industry's reactor pressure-vessel integrity program. It is essential that a nuclear reactor be operated in a manner that ensures that the vessel integrity is maintained under both normal and transient operating conditions. In particular, the vessel must be protected from both brittle and ductile fracture. This is accomplished by postulating limiting flaws and using linear elastic fracture mechanics (LEFM) models to calculate the allowable coolant temperature (T) and pressure (P) during heat-up, cool-down, and leak/hydro testing (P-T curves). The P-T limits are revised periodically throughout the life of a plant to account for neutron damage to the pressure vessel. The Charpy shift, indexed at 41 J, is used to shift the ASME (American Society of Mechanical Engineers) reference stress intensity factor (K_{IR}) curve to account for the effects of neutron bombardment.

Since it is impractical to test large fracture toughness specimens throughout the life of a nuclear power plant, surveillance programs use Charpy and tensile specimens to track the neutron-induced embrittlement. As illustrated in Figure 6, the nuclear industry uses the 41 J index to define a ductile-brittle transition temperature (DBTT). The effect of neutron irradiation is to shift the transition region to higher temperatures (ΔT_{41}) and the Nuclear Regulatory Commission (NRC) sets screening limits on the maximum shift in the energy-temperature curve that can occur during the life of the plant. If the screening limits are exceeded, then either the plant must be shut down or a thermal anneal must be conducted to restore the material properties.

The ability of the material to withstand ductile fracture is judged by the upper shelf energy (USE). In older plants built before fracture toughness testing was widely used, Charpy testing was used to qualify individual heats of material. The ASME code and the Code of Federal Regulations prescribe minimum plate properties that must be satisfied prior to service (e.g., at least 102 J of energy on the upper shelf prior to service). The NRC requires an in-depth fracture-mechanics assessment if the Charpy USE is expected to drop below 68 J during the operating life of the plant.

It is difficult to include all the economic benefits realized by using the Charpy impact test in the nuclear industry. Nevertheless, some insight can be gained by noting that most utilities assess the outage cost and loss of revenue for a nuclear plant in the range of $300,000 to $500,000 per day. If Charpy data can be used to extend the life of a plant one year beyond the initial design life, a plant owner could realize revenues as large as $150,000,000. Further, the cost avoidance from a vessel-related fracture is expected to be the billion-dollar range. To date, the NRC has shut down one U.S. plant as a result of Charpy data trends. It is important to note that this plant's pressure vessel was constructed from a one-of-a-kind steel and is not representative of the US reactor fleet. Nonetheless, with decisions like this being based on the Charpy test, the importance of ASTM E 23 and the restraints it applies cannot be overemphasized

Cost Saving in the Steel Industry

The Charpy V-notch (CVN) test specimen and associated test procedure is an effective cost-saving tool for the steel industry. The specimen is relatively easy to prepare, many specimens can be prepared at one time, various specimen orientations can be tested, and relatively low-cost equipment is needed to test the specimen. In many structural steel applications, the CVN test can be used (1) as a quality-control tool to differentiate heats of the same type of steel, (2) for quality assurance purposes, and (3) to predict service performance of components. Also, CVN test information can be correlated with fracture toughness data for a class of steels so that fracture mechanics analysis can be applied directly. One may question how all the above factors help a steel producer sell a reliable product

Consider the following case. A steel producer has a contract to supply plate steel for an off-shore platform. The plate material needs to meet mechanical properties that are quite rigorous for reasons of safety and end-product reliability. Before full-scale production of the order can begin, the steel supplier needs to demonstrate to the buyer that the material is capable of meeting such criteria. To accomplish this, the supplier qualifies the material for the project. The process begins by making the steel grade and then testing a portion of the plate to determine whether all required criteria are met. Steel-mill equipment imposes limitations on plate size; therefore, individual steel plates need to be welded together in the field to produce lengths that can reach deep into ocean waters. Small sections of the sample plate are welded together, and fracture-mechanics tests are conducted to determine the crack tip's opening displacement (CTOD) toughness in the heat-affected zone (HAZ) and areas and along the fusion line where the weld metal meets the base metal. Then, for example, a steel supplier might correlate the CTOD test results with CVN 50 percent ductile-brittle transition temperature (DBTT). By agreement

between the customer and supplier, this correlation can allow the steel supplier to use the Charpy test instead of the more expensive and time-consuming CTOD testing.

One piece of information that can be used directly in design applications is the critical plane-strain stress-intensity, K_{Ic}, value. It is defined as that value occurring ahead of a sharp crack at the moment of unstable crack propagation. The K_{Ic} value is related to component geometry, applied stress, and flaw size. Barsom and Rolfe [15] proposed a plane-strain fracture toughness CVN energy correlation that can be used in the transition region:

$$K_{Ic}^2 / E = 2(CVN)^{3/2} \quad \text{(English units)},$$

where K_{ic} is the plane strain fracture toughness, E is Young's modulus, and CVN is the absorbed energy value from the Charpy V-notch test.

If one knows the values of E and CVN (easily obtainable) for a given material, a quantitative assessment of permissible stress levels and critical flaw size can be calculated. It must be noted that the above equation was developed for a particular grade of steel and therefore may not be suitable for all grades of steel. However, the development of such correlations for a particular class of steel can be very cost effective. As mentioned earlier, the nuclear industry has used Charpy parameters qualitatively to indicate the need for in-depth fracture mechanics analyses when the Charpy parameter falls below prescribed values. Design criteria for bridge steels have also been based on such correlative procedures.[16]

Continuing Standardization Efforts

Even after 100 years, the Charpy impact test procedures still have room for improvement. The ASTM E23 standard has recently been redrafted to provide better organization and to include new methods such as in-situ heating and cooling of the test specimens. Two new related standards are also under development through ASTM Task Group E28.07.08 on Miniature and Instrumented Notched Bar Testing, which was formed a little more than two years ago. The first standard covers miniature notched-bar impact testing and relies on many of the existing practices related to test machine requirements and verification as specified in existing standard E 23. The second standard is focused on instrumented testing. This involves the use of strain gages that are attached to the striker. In this method, the force-deflection curve can be obtained for each test. Research is focused on using these data to obtain plane-strain fracture toughness as well as other key test parameters. Upon acceptance of the standard by ASTM, both the existing E23 standard and the new miniature notched-bar standards would reference the instrumented impact standard.

Conclusion

As can be seen, the ASTM E 23 standard is a document that is improving with increasing technical knowledge. Several years ago, at the ASTM Symposium on the Charpy Impact Test: Factors and Variables, a bystander was overheard to say, "I see that there is a Symposium on the Charpy Test; what can be new there?"[17] Since then, the document has been updated twice and is currently being revised to reflect new developments and to make it more "user friendly".

Although Method E23 is now a standard, and no longer tentative, it continues to be a "work in progress;" a work used extensively to help evaluate existing and new materials for products and structures - a test to insure safety as well as reduce the initial and life-of-structure costs. Knowledge is continually being gained that will help make the test more accurate and reliable. New technologies such as miniaturization of the test, instrumenting the striker to obtain additional data, and developing mechanics models to enable extraction of plane strain fracture toughness will be areas of development over the next 100 years. It is anticipated that the improvements to E 23 obtained over the past 100 years will not overshadow the benefits that will be realized in the future.

References

1. Louis Révillon, "The definition of Resilience in Impact Tests," Proceedings of the Vth Congress of the Internaional Association for testing Materials, Section III3, Copenhagen 1909.
2. Tredgold T., "Strength of Cast Iron", pp 245 - 268, 1824.
3. White A.E., Clark C.L., "Bibliography of Impact Testing", Department of Engineering Research, University of Michigan, 1925.
4. LeChatalier A., "On the Fragility After Immersion in a Cold Fluid", French Testing Commission, Volume 3, 1892.
5. Experiments with a New Machine for Testing Materials by Impact, S.B. Russell, Transactions ASCE, Vol 39, p. 237, June 1898.
6. Bibliography on Impact Tests and Impact Testing Machines, W.K. Hatt and E. Marburg, Proceedings ASTM, Vol. 2, p. 283, 1902.
7. Report on the Work of the Council from the Budapest to the Brussels Congress - 1901-1906, Proceedings of the International Association for Testing Materials, Brussels Congress, 1906.
8. Proceedings of the Sixth Congress of the International Association for Testing Materials, New York, 1912.
9. Report on Impact Tests of Metals, G. Charpy, Proceedings of the Vth Congress of the International Association for Testing Materials, Section III_1., Copenhagen 1909.
10. The Design and Methods of Construction of Welded Steel Merchant Vessels: Final Report of a (U.S. Navy) Board of Investigation, Welding Journal, 26, 7, July 1947, p. 569.
11. N. H. Fahey, Effects of Variables in Charpy Impact Testing, Materials Research Standards, Vol 1, No. 11, Nov., 1961.
12. D. E. Driscoll, Reproducibility of Charpy Impact Test, Impact Testing, ASTM STP 176, 1955.
13. N. H. Fahey, The Charpy Impact Test - Its Accuracy and Factors Affecting Test Results, Impact Testing of Metals, ASTM STP 466, ASTM, 1970.
14. C. N. McCowan, C. M. Wang, and D. P. Vigliotti, Summary of Charpy Impact Verification Data: 1994 - 1996, Submitted to the Journal of Testing and Evaluation, 1998.
15. J. M. Barsom and S. T. Rolfe, Eng. Frac. Mech., 2(4), 1971, p. 341.
16. ASM Metals Handbook, Ninth Edition, Vol. 8, 1985, pp. 263, 264.
17. Charpy Impact Test: Factors and Variables, ASTM STP 1072, J. M. Holt, editor, ASTM, 1990.

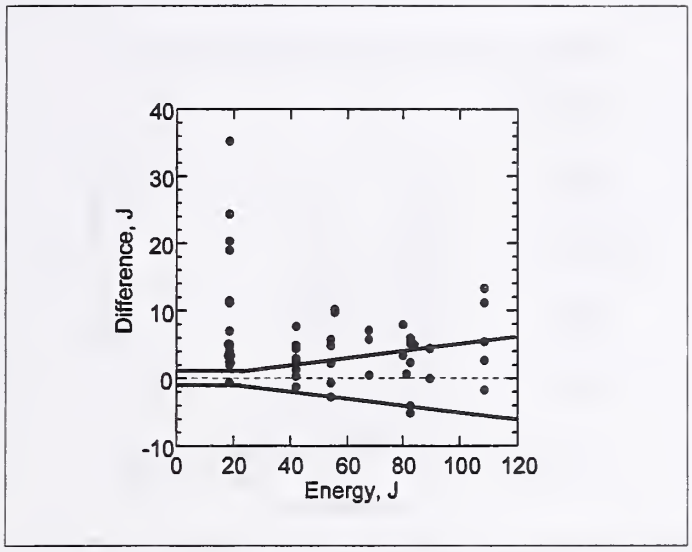


Figure 1: The deviation and energy values obtained for the first round of tests on industrial machines. The deviation is calculated as the difference between the results of the Watertown Arsenal's machines and the industrial machines. These data were originally published by D. E. Driscoll, Reproducibility of Charpy Impact Test, ASTM STP 176, 1955.

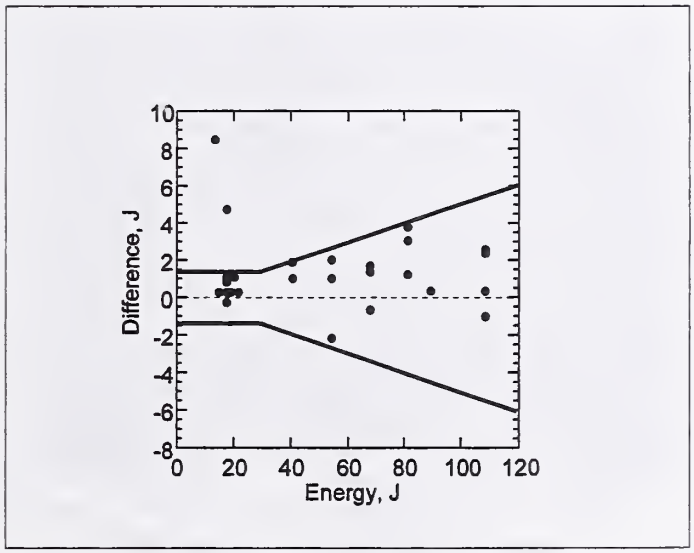


Figure 2: The deviation and energy values for the second and third rounds of tests on industrial machines. The data show all but two of the machines tested were able to pass the 1.4 J or 5% criteria after appropriate repairs were made. These data were originally published by D. E. Driscoll.

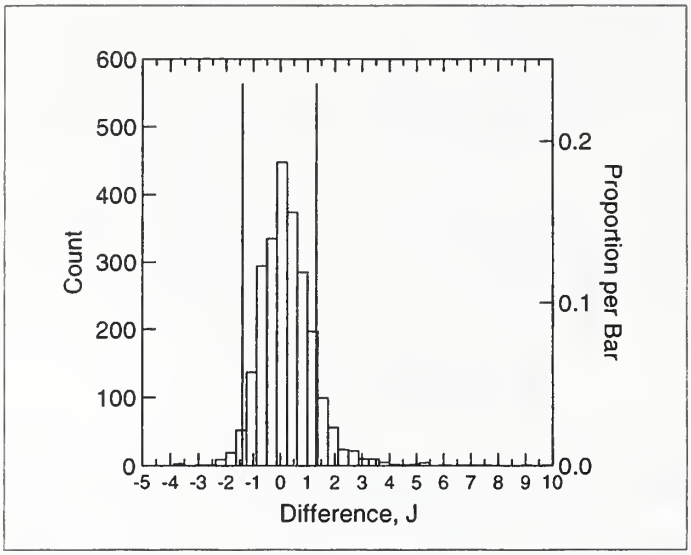


Figure 3: Distribution of low-energy verification data. Data for 1995-1997. Approximately 2,400 tests; each test is an average for five specimens.

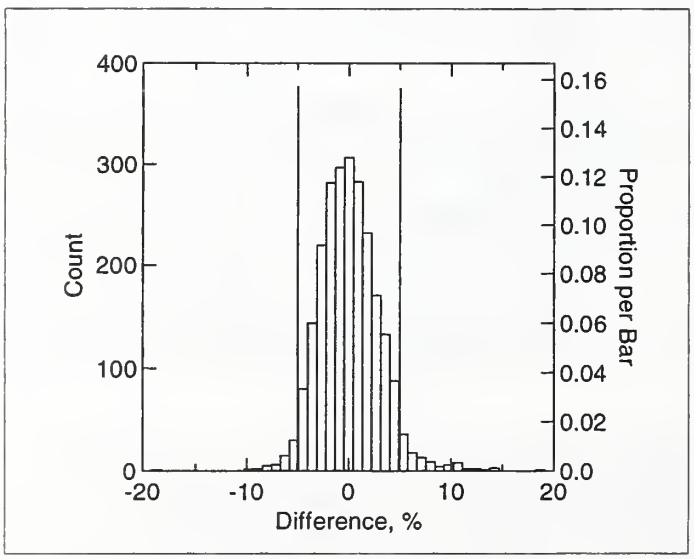


Figure 4: Distribution of high-energy verification data. Data for 1995-1997. Approximately 2,400 tests. Each test is an average for five specimens.

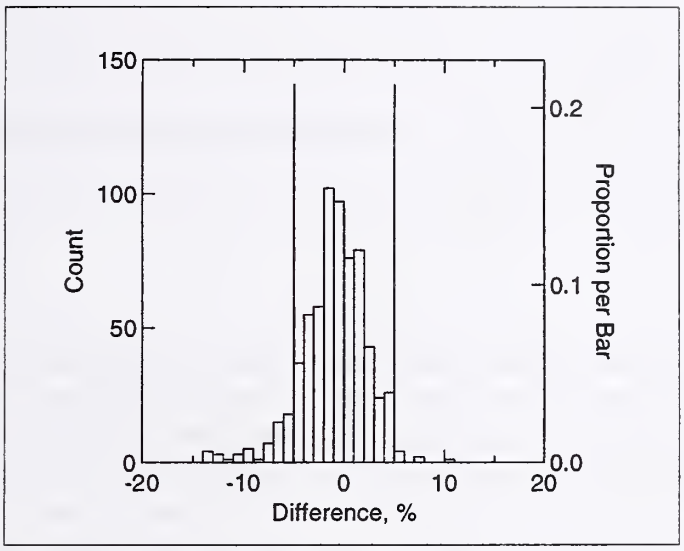


Figure 5: Distribution of the super-high energy verification data. Data for 1995-1997. Approximately 650 tests. Each test is an average for five specimens.

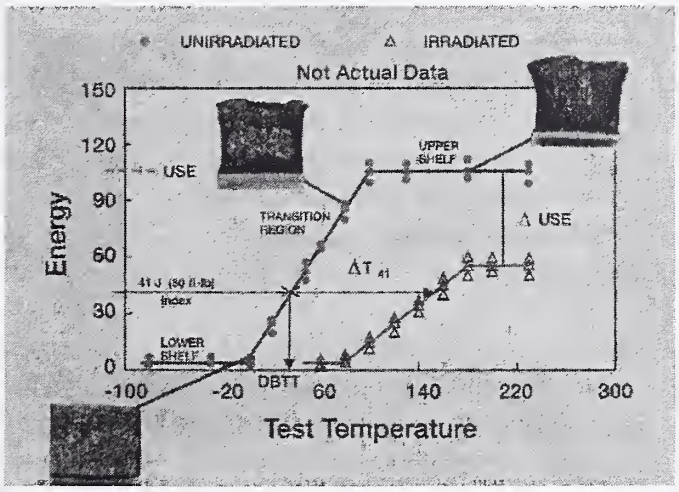


Figure 6: Some materials, such as ferritic pressure vessel steels, exhibit a transition in fracture behavior as the notched-bar impact test temperature is increased. At low temperatures the fracture is predominantly cleavage; at intermediate temperatures the fracture is a mixture of both cleavage and ductile; and above the transition region the fracture is entirely ductile.

T.A. Siewert[1]

Impact Test Methods: Procedures and Analysis

Abstract

This report discusses recent research in the development and refinement of impact test procedures, as well as some direct and indirect verification procedures to assure that the machine is operating properly. Most of the recent research can be organized into several broad categories: the specimen (e.g., surface finish, tolerances, and miniature sizes), the anvils and striker (e.g., radii and surface finish), and general test procedures (e.g., time to reach test temperature and suitability for cryogenic testing). This report also provides some direct and indirect verification tests that can be used to evaluate the mounting and condition of the anvils and striker on the machine. Finally, this report concludes that an open, international discussion of the procedural details of the test is one of the best ways of promoting harmonization.

Keywords

impact testing; international intercomparison; machine verification; specimen notching and conditioning; striker radius; test procedures; test temperatures

Introduction

For over 100 years, researchers have been trying to understand and to measure the effect of impact loading on the performance of engineering materials. In fact, the development of consistent impact procedures was recognized to be of such importance that, even in 1912, Committee 26 (on impact testing) of the International Association for Testing Materials (IATM) summarized its main goals as to "fix the conditions to be fulfilled by two distinct tests in order that the results may be comparable and to correlate these numerically definite results to the qualities determining the practical values of a material for different uses"[1].

Since then, impact-test procedures and analytical methods have been refined as various researchers have discovered additional parameters that affect the test results. In some cases, these new results have been widely and uniformly adopted. In other cases, different standards organizations or machine manufacturers have chosen different approaches. As a result of many such choices by the different standards organizations over the years, we now find some variation in impact test procedures around the world. Certainly, worldwide comparison of test data would be simplified if the procedures could be further harmonized between countries and between the various standards. The following section describes recent work directed toward understanding

[1]Materials Reliability Division, National Institute of Standards and Technology, Boulder, Colorado 80305 USA

the effect of various procedural details. Publication of such work can persuade the various standards committees around the world to choose the best procedural details (that produce the most consistent results) or to determine that some existing differences in procedures have no effect (so data developed under different procedures are considered equivalent and are mutually recognized).

Recent research on procedures

The three most common impact test procedures in use around the world are probably ISO R 442 "Metallic Materials – Impact Testing – Verification of Pendulum Impact Machines," ASTM E 23 "Standard Test Methods for Notched Bar Impact Testing of Metallic Materials," and JIS Z2242 "Method for Impact Test for Metallic Materials". While the three have some similarities, they also differ. Much current research is directed toward both improving these (and other) standardized procedures, and trying to understand the effect of their differences and move toward harmonization.

Striker Radius

Perhaps the aspect of the procedure that has been receiving the most study recently is that of striker radius. At least 5 papers in the past 12 years have investigated the differences between using strikers of 2 and 8 mm radius.

In 1989, Fink [2] compared the results of a number of variables, including notch preparation and striker radii, on impact data. He studied a number of types of steel including 4340 (a heat-treatable tool steel used for 15 and 100 J reference specimens), ASTM A 537 (a carbon-manganese steel used in pressure vessels), and HY-80 (a quenched and tempered high-strength steel). For these steels, he reported a nearly 1:1 relation for data generated with the two strikers, where the 2 mm striker produced results about 4 % higher. The precise relationship was:

$$E_2 = (1.042) * E_8 + 0.516, \qquad (1)$$

where E_2 and E_8 are the energies of the 2 mm and 8 mm strikers and where the values are in ft-lbf. He reported a coefficient of determination (r^2) of 0.9987 and a standard error of 1.36.

Also in 1989, Naniwa et al. [3] compared the results of impact machines with strikers of 2 and 8 mm radius using steels with a range of absorbed energies. Figure 3 in their report showed a 1:1 trend for the two striker radii up to about 200 J, then a gradual increase in the data of the 8 mm radius striker. When the machine with the 2 mm striker reported 300 J, the machine with the 8 mm striker reported about 400 J. Data for percent shear and lateral expansion showed no trend, and the shape of the transition curves were the same. Additional impact testing with instrumented strikers showed that both striker designs produced similar shapes for the first parts of the record, but the load was substantially higher near the end of the record for the 8 mm radius striker. This suggested that the difference was occurring near the end of the loading cycle. Further testing (load measurements combined with imaging of the specimen in the anvils) on a static bending machine confirmed that the higher energy in the tail region occurred when the

sides of the impact specimen made contact with the shoulder of the 8 mm striker. Thus, they attributed the difference between the two strikers solely to the wider width of the striker with 8 mm radius.

In 1995, Nanstad and Sokolov [4] evaluated the data from machines with the 2 mm and 8 mm strikers using six different materials. They studied two heats of ASTM A 533 (a pressure vessel steel), a submerged arc weld with a high upper-shelf energy, a submerged arc weld with a low upper-shelf energy, a Russian Cr-Mo-V forging steel, and two kinds of reference materials (4340 steel and a maraging steel). Although one plate showed lower values with the 8 mm striker and the other plate (and the low upper-shelf weld) showed lower values with the 2 mm striker, they concluded there was no clear trend (within one standard deviation) up to 175 J. Only the highest-energy reference material (near 220 J) showed a clear difference, where the 8 mm striker produced energies about 11 % higher than those for the 2 mm striker.

Also in 1995, Siewert and Vigliotti [5] evaluated the data from two different brands of U-type pendulum machines, each with both the 2 mm and 8 mm radius striker. They used reference grade specimens at energies of 18, 45, 100, and 200 J. The small standard deviation produced by these specimens allowed a very precise measurement of machine or striker effects. They found very small differences between the two strikers for the three lower energy ranges, and an even smaller effect between the two brands of machines. At 200 J, they noted that the 8 mm striker produced energies about 10 J higher than those for the 2 mm striker, and the 8 mm striker produced standard deviations that were higher by about a factor of three.

These four studies suggest that the striker radius does not seem to be an important variable up to about 150 or 175 J, at least for the steels that were evaluated. However, above 200 J the two striker radii produce divergent results as the energy increases. Ruth [6] has attempted to produce a compromise striker, one with the narrower profile of the 2 mm striker, but also with a flatter surface on the very front edge. This was accomplished by grinding the front to an 8 mm radius, then blending this surface to the edges by the use of a 1.5 mm radius. So far, this approach has not reached its goal.

Specimen Fabrication

Koester and Barcus [7] compared grinding and broaching of the notch. They found that both procedures produced data that were equally consistent, but there was a bias between the two techniques. They attributed this to differences in either the microstructural damage due to the notching or to slight imperfections in the notch radii.

Fink [2] also looked at notch production by grinding, broaching, and milling (with a fly cutter). He concluded that the grinding of the notch produces the smoothest and most consistent profile.

Direct Verification (Machine Condition and Mounting)

Schmieder [8] found that direct verification of a machine is not a simple task. He based much of his work on the concept that the permitted inaccuracy of a metrological measurement must be 10

times better than the tolerances specified for the device. In other words, he tried to use instruments and techniques that were more precise by an order of magnitude than what was required by the standard, to develop a better estimate of how close the machine was to exceeding the prescribed tolerance.

He evaluated four C-type pendulum machines and five U-type pendulum machines, spanning capacities from 3 to 2500 J. He found that the losses due to friction and windage could exceed the permitted limits on machines of very small capacity or on multi-range machines (where the bearings are sized for the highest capacity, and so have too much friction for the lower capacity). He also found that checking the difference between the center of strike and the center of percussion requires extremely accurate measurement of the period of the pendulum (as the center of percussion varies as the square of the period of the pendulum). At a 5 degree angle of swing, the friction would damp the swing before enough cycles would occur. At higher angles, the nonlinear terms became important, and even the use of elliptic integrals in the theoretical analysis was unable to correct for these effects.

Schmieder et al. [9] later studied the effect of various machine dimensions, including: tilted anvils, thinner anvils (striker contacting anvils 5 mm past the normal position), and striker not contacting the specimen opposite the notch. All these were studied at levels in excess of the variation permitted by ASTM Standard E 23, and all variations noticeably increased the absorbed energies. Thus, these data support keeping the machine tolerances that are specified by E 23.

Porro et al. [10] studied the use of compliance to evaluate the quality of the machine mounting, in terms of such common problems as loose bolts on the base of anvils, paint or other low-friction materials under the base.

Ruth et al. [11] studied the effect of surface finish of the machine anvils and striker. They found that surfaces smoother than those required by the standard procedures better simulate the surfaces of these parts after a period of use. Thus, a better finish will reduce the discontinuity in apparent energies when these parts are replaced.

Ruth [11] also studied the effect of radius on the corners of the 8 mm striker, because a 0.25 mm tolerance can rapidly wear beyond the tolerance. He found that increasing this radius to 0.5 mm had little or no effect, but increasing the radius to 1 mm had a very measurable effect.

Yamaguchi et al. [12] studied the effect of anvil radius and taper. They reported a measurable reduction in absorbed energy as the taper angle increases from 9 to 12 °, and a 5 % change in energy as the anvil radius increases from 1 to 1.5 mm.

Effects of Specimen Size and Dimension

Alexander and Klueh [13] compared Charpy specimens of standard size (10 mm by 10 mm) to specimens of half and third size. They found that the smaller specimens allowed more specimens to be produced from a given amount of material (especially important for irradiated material), but produced different upper-shelf energies and different transition temperatures. They concluded

that the upper-shelf energies could be corrected with a simple volumetric normalization procedure, but the shift in transition-temperature was more complex. Later, Alexander et al. [14] revisited this issue and developed sub-size verification specimens that could be used to verify the performance of machines used to test sub-size specimens.

Manahan et al. [15] also looked at sub-size specimens, and developed a test machine design. They proposed a minimum cross section of 5 mm by 5 mm, and recommended side grooves to increase the amount of material in these smaller sections that is exposed to plane strain conditions.

Marsh [16] studied the effect of changing the tolerance on the right angle between the two 10 mm faces of the specimen. He varied the angles outside of the tolerance of 10 minutes of arc and found that greater variations produced statistically significant changes in the energies, especially for specimens with absorbed energies near 100 J. He concluded that a tolerance of 10 minutes on the right angles should be maintained.

Test Temperatures and Specimen Conditioning

Nanstad et al. [17] studied the effect of thermal conditioning, the process of bringing the specimen to the desired test temperature. They investigated a number of media including water, oil, acetone, and methanol, at temperatures above and below ambient. They found that water was a poor choice between 50 and 100 °C because evaporative cooling is so significant that the specimen may cool below the temperature tolerance even if the specimen is broken within 5 s of leaving the bath. Also, they found that soaking times used with gaseous media need to be increased to assure that the specimen has reached thermal equilibrium.

The growing use of cryogenic magnets has promoted the use of impact testing to measure the ductile-brittle transition of structural materials at temperatures down to 4 K. Tobler et al. [18] offer several cautions. They found that the very low specific heat of metals below 77 K causes the specimens to heat rapidly as they are transferred from the bath to the anvils. For this reason alone, valid tests cannot be performed according to the procedures of E 23. Further, even cooling the specimen in place in the anvils is unable to provide accurate data, as the work hardening during the initiation and propagation of the crack raises the temperature substantially. Thus they concluded that pendulum impact testing is not valid below 77 K, and any attempt to correlate performance at 4 K from specimens cooled to 4 K is confounded by the variations in work-hardening rates in the various materials.

Manahan [19] reported that conditioning of the specimen when on the machine anvil and in position for testing (by use of a special fixture) reduces the changes in temperature that can occur when a specimen is transferred from a conditioning bath to the anvils. In addition, this procedure doubles the precision in centering the specimen in relation to the striker, since there is no rush to position the specimen.

Other Procedure Details

Sundqvist and Chai [20] reported on the production of in-house standard specimens (from a stable nickel-based alloy) for tracking the performance of an impact machine between the formal reverifications required by standards. They found that this was an excellent method of tracking the performance of machines that are used to test specimens that lead to excessive wear of the striker and anvils.

In spite of the widespread use of notched specimens for evaluating material performance, Galban et al. [21] report that unnotched specimens can provide standard deviations as small as, or smaller than notched specimens of the same material. Since verification of machine performance is separate from evaluation of material performance, use of such specimens (with low standard deviations) could reduce the cost of the verification specimens.

Comparison of Data – Machine-to-machine and country-to-country

Several recent round-robins or comparisons of national reference machines confirm that today's Charpy test procedures are at least as reproducible as those reported by Driscoll [22] in 1955, and are consistent between countries. These recent round robins have shown that the certified energies of verification specimens distributed by national metrological authorities usually agree within 1 % with the values determined by other national authorities. A 1998 study [23] compared the four organizations or laboratories that were found to certify the verification specimens for Charpy impact machines. These organizations were the Institute for Reference Machines and Measurements (IRMM, in Belgium), Laboratoire National D'Essais (LNE, in France), National Institute for Standards and Technology (NIST, in the United States), and National Research Laboratory for Metals (NRLM, in Japan). The study involved a comparison of the 2 and 8 mm radius strikers, three absorbed energy levels, and a large number of replicate tests for each of these conditions at each of these organizations. This study concluded that the other organizations developed average energies very close to those assigned by the laboratory that produced them, the specimens produced by the four organizations have similar spread in the data (coefficient of variations between 0.02 and 0.04), and the 2 and 8 mm radius strikers produced similar results for 4340 steel (absorbed energies below 200 J). Therefore, in spite of the various differences in procedures between the major standards in use around the world, the basic test procedure is quite reproducible, so the results developed in different countries and on different designs of (verified, high-quality) machines can be compared with confidence.

A follow-on three-year study [24] is currently underway. This study is examining the effect of the differences in the measurement systems (e.g., master machines versus master batches) used by various National Measurement Institutes (NMIs) that distribute reference specimens. Ten

specimens at each of three energy levels (15 J, 30 J, and 110 J) will be tested on seven reference-grade machines every six months. The data are expected to allow comparison of the verification systems used by the laboratories, as well as provide data on machine variables, offsets, and uncertainties relative to the harmonization of the respective systems.

Machine Installation

The data from a machine are not reliable unless the machine is mounted properly. The following is the detailed procedure used by NIST to mount its three Master Charpy Reference Machines. Unless the manufacturer of your machine supplies alternate instructions, you should consider using these guidelines. These are taken from a NIST Technical Note 1500-8 [25].

A stable foundation for the impact machine is critical to ensure accurate results. Energy losses through the foundation must be kept to a minimum. The recommended foundation is a high-strength concrete that measures about 1.5 m long by 1 m wide by 0.5 m thick. Usually you will need to cut a hole in the floor to accommodate the new foundation. If other equipment in the area could affect the machine operation, you may want to isolate it from the floor with expansion-joint material.

Hold-down bolts used to secure the machine to the foundation should be of the inverted "T" or "J" type. (The next section, on direct verification, describes problems with the use of lag bolts, which may be tightened up against the base without gripping the concrete.) The bolts, nuts, and washers should have a high strength (for example, AISI grade 8 or higher). NIST recommends using bolts or rod with a diameter of 22 mm. NIST used 22 mm diameter grade 8 threaded rod, cut into pieces that were about 600 mm long. Then, 150 mm long pieces of the same threaded rod were welded to the end of the 610 mm (24 in) pieces to make inverted "T" bolts.

The machine was positioned over the center of the foundation hole. The machine was held approximately 100 mm above the floor by using spacers suitable to hold the weight of the machine. The "T" bolts were positioned in the machine-base mounting holes with a nut below and above the base of the machine. The nuts were tightened to keep the "T" bolts straight while the concrete was poured. The ends of the T bolts were positioned approximately 25 mm from the bottom of the hole. The machine was then leveled on the spacers. Leveling did not need to be as accurate as the final leveling (mentioned later). Reinforcement bars were attached to the top of the horizontal rod previously welded to the bottom of the "T" bolts. The reinforcement bars were attached in the form of a box connecting the four bolts. A second box formation of reinforcement rods was attached to the "T" bolts 0.25 m above the first box. The concrete was then poured under the machine. The concrete was finished as level as possible at this time. Before the concrete fully hardened, we removed enough concrete, to a depth of approximately 25 mm, from around each "T" bolt to be able to thread a nut to below the surface of the concrete, and cleaned the exposed threads. The machine was left in this position for 72 hours.

After 72 hours, the nuts on top of the base plate were removed and the machine was lifted off the "T" bolts. The bottom nuts were then threaded down into the cavities that had been created before the concrete hardened. The nuts were left high enough on the "T" bolts to enable the use of an open-end wrench to adjust them after the machine was positioned on them. At this point, the base of the machine was lightly coated with oil to keep the grout from adhering to it. The machine was then lifted back onto the "T" bolts and was positioned on the adjustment nuts. The machine was now ready to be leveled. A machinist's level was used to ensure meeting the level tolerance of 3:1000. The critical leveling procedure was done using the four nuts under the

machine. After the machine was leveled, the outsides of the nuts were wrapped with duct-seal putty to facilitate their removal from the "T" bolts later in the process.

At this point the base of the machine was ready to grout. Heavy cardboard forms were placed around the base of the machine to keep the grout under the machine. The grout was pushed under the machine, so that the base of the machine was in total contact with the grout. The machine was left in this position for 72 hours.

After 72 hours, the machine was lifted off the "T" bolts one last time. The grout was inspected for cavities and for surface contact with the bottom of the machine. The putty was removed from around the nuts. Some grout leaked around the putty and had to be chipped away from the nuts to enable them to turn. The supporting nuts were removed from the "T" bolts. After all debris was removed from the grout, the machine was repositioned on the "T" bolts and rested on the grout. Washers and nuts were installed on the "T" bolts and were tightened to pull the machine tightly against the grout. The level was checked at this point. The "T" bolts were cut off to approximately 12 mm above the nuts. The nuts were torqued to about 500 N-m. The level of the machine was rechecked at this point.

Direct Verification

This section explains direct verification requirements (based on those in ASTM Standard E 23), that confirm that a machine is in good operating condition, without the use of verification specimens. The direct verification tests are physics-based tests, which assure that the machine is functioning as closely as possible to the behavior of a simple pendulum, with only small losses due to friction and windage. Direct verification is most important when the machine is first installed or when major parts are replaced, but is also important during the periodic reinspections. While these tests are required for the periodic reinspection in ASTM E 23, NIST recommends that the free-swing test and windage-and-friction test be performed each day that the machine is used. The records of these tests then serve as a convenient measure of bearing performance. The following recommendations also come from a NIST Technical Note [24]. Space limitations prevented including illustrations of these characteristics here. The illustrations are available in the Technical Note.

Since the Charpy test is a dynamic test with vibration and impact loads, the hold-down bolts may loosen over time. In extreme cases, this may introduce error sufficient to cause a machine to exceed the tolerance limits of the indirect verification test. In marginal cases, the movement may still be sufficient to add a bias to the results that reduces the likelihood of passing. The tightness of all bolts should be checked periodically, especially the anvil bolts, the striker bolts, and the base-plate bolts. The manufacturer can supply the torque values for the anvil and striker bolts. The base-plate bolts should be torqued to the recommended torque values for the grade and size of the nuts and bolts. Only “J” or “T” bolts should be used; lag-type bolts can lead to errors. These are made to withstand only static loads. We believe that in some cases, the insert portion of lag bolts can loosen in the concrete. When lag bolts are retightened, they can pull out of the concrete and be pulled against the base of the machine, giving the impression of a properly mounted machine. This condition is very difficult to detect. A machine with this problem will

exhibit erroneously high energy values at the low-energy level. The mounting procedure used to eliminate this problem for the NIST Master Reference Machines was described in the previous section.

Standard E 23 describes a routine check procedure that should be performed weekly. It consists of a free-swing check and a friction-and-windage check. The free swing is a quick and simple test to determine whether the dial or readout is performing accurately. A proper zero reading after one swing from the latched position is required on a machine that is equipped with a compensated dial. Some machines are equipped with a non-compensated dial. Such a dial is one on which the indicator cannot be adjusted to read zero after one free swing. The user should understand the procedure for dealing with a non-compensated dial. This information should be available from the manufacturer.

The friction-and-windage test assesses the condition of the bearings. The pendulum should be released and allowed to swing 10 half cycles (5 full swings). (The release mechanism should be held down this whole time to avoid additional friction when the pendulum swings back up to where it may push on the latch.) As the pendulum starts its eleventh half swing, the pointer should be reset to about 5 % of the scale capacity. Record the actual value and divide by the eleventh half swings. Divide this number by the machine range capacity, then multiply by 100. Any loss of more than 0.4 % of the machine capacity is excessive, and the bearings should be inspected.

Keeping a daily log or shift log with the machine is also recommended. The log can be used to track the zero and friction values. The log can also include information such as number of tests, materials tested, maintenance, and any other useful comments.

The anvil and striker radii should be carefully inspected for damage and for proper dimensions. Damage (chips or burrs) can be detected easily by visual inspection and by running a finger over the radii to check for smoothness. Measurement of the dimensions requires more sophisticated equipment. Radius gages are usually inadequate to measure the critical radii. Making molds of the radii (such as with silicone rubber) or making an indentation in a soft, ductile material (such as annealed aluminum), then measuring the impressions on an optical comparator is recommended. Occasionally even a new set of anvils and striker may have incorrect radii. Thus, new anvils and strikers should always be inspected before being installed in the machine. Since the radii will not have local wear before use (the radii are consistent along their length), they can be measured directly on an optical comparator or other optical measurement system.

Indirect Verification

Indirect verification uses carefully characterized test specimens to stress the test machine components to levels similar to those experienced during routine usage. Since many machine problems, such as loose anvils or striker, cannot be detected during direct verification, indirect verification serves as an important supplemental test of the machine performance.

Centering tongs, such as those described in ASTM Standard E 23, are an excellent way to insert the specimens at the very center of the anvils. The tongs should be inspected occasionally for wear or damage. A proper set of tongs is critical for the accurate placement of the specimen. Some machines are equipped with a centering device. The device should also be inspected for wear and proper operation. Centering devices for low-temperature testing should be evaluated carefully because the centering operation can extend the time between a specimen's removal from the bath and fracture, and so may exceed the five-second interval allowed for transferring and fracturing the specimen.

Some reference specimens are designed to be tested at −40 °C (−40 °F). Since the absorbed energy changes with temperature, accurate temperature control is necessary to obtain valid test data. The temperature indicator should be calibrated immediately before testing. Ice water and dry ice are very convenient calibration media.

Post-Fracture Examination

Just matching the reference energies is not sufficient to confirm that the machine is fully satisfactory. For example, worn anvils can combine with high-friction bearings to compensate for each other and produce an artificially correct value during the verification test. These are called compensating errors. Unfortunately, these errors compensate only over part of the range, so the machine produces generally inaccurate values. The post-fracture examination of standardized verification specimens is a good way to identify such effects. Therefore, the NIST verification specimens come with a questionnaire (with critical questions about the machine and the test procedure) and a mailing label so the specimens can be returned to NIST. All specimens are examined and compared to the data on the questionnaire before a response is sent to the customers.

Following are the most common of the problems observed during examination of fractured specimens. In many cases, suggestions on how to correct or avoid them in the future are included.

Worn Anvils

Most of the wear of an impact test machine occurs on the anvils and striker. This wear can be evaluated by examining the gouge marks that are formed on the sides of high-energy specimens when they are forced through the anvils. Anvils that are within the required tolerance of the standard will make a thin, even gouge mark all the way across both pieces of the broken specimen. As the anvils wear, they will make a wider, smeared mark across the specimen halves. When wide, smeared marks are observed on a customer's specimens, the anvils should be changed, because the reduction in energy needed to push the specimens through worn anvils eventually drops the machine below the lower tolerance in the energy range. You can monitor the wear on your machine by retaining some specimens that are tested with new anvils and comparing them to specimens of similar composition and hardness that are tested as the anvils wear. For specimens at a similar absorbed energy, the gouge marks will grow wider and smoother as the anvils wear.

Off-Center Specimen

An off-center specimen strike occurs when a specimen is not centered against the anvils, so the striker contacts the specimen to the side of the notch. The low-energy specimen best indicates when an off-center strike occurs. This condition can be identified on the specimens by finding that the gouge marks caused by the anvils are not equidistant from the machined notch edges, and the striker gouge mark is offset the same amount from the notch. Also, the fracture surface of a correctly tested low-energy specimen is flat and both halves are even. However, the fracture surfaces of a specimen that has been tested off-center are on an angle. The more off-center the strike, the steeper the angle will be. This problem increases the energy needed to fracture a specimen. The most common causes for this slipping are worn or damaged centering tongs, a worn or misaligned machine centering device, careless test procedures, or the use of a cooling fluid that is too viscous at the test temperature, which causes the specimen to float on the specimen supports. Most machine manufacturers should be able to provide new centering tongs. Ethyl alcohol seems to be one of the best cooling medium to prevent floating because it evaporates quickly from the bottom of the specimen.

Off -Center Striker

This differs from the off-center specimen in that the specimen is centered against the anvils so the anvil gouge marks are equidistant from the machined notch edges. However, the striker does not contact the specimen precisely opposite the notch. An off-center striker is usually attributed to the pendulum shaft shifting off center. This shift can be the result of a loose alignment ring on the shaft or a loose bearing block on the machine. This problem also increases the energy needed to fracture specimens at all energy levels.

Uneven Anvil Marks

Frequent testing of subsize specimens can cause the anvils to wear unevenly. Since this wear is restricted to only a fraction of the area that the full-size reference specimen contacts, there is usually no effect on the energy required to fracture the specimen. This anvil condition presents two problems. First, since subsize wear is usually not indicated by a change in the energy required to break a reference specimen, inspection of the broken specimen is required. This wear will cause the anvils to be out of tolerance according to the requirements in the standard. This means that the machine does not meet the direct verification requirements of the standard and is therefore, not eligible for the indirect verification process. The second, and more important problem, is that the subsize specimens are being tested in an area of the anvil that is worn. When the wear is substantial, this condition will produce artificially low subsize values of energy. The anvils on a machine with this condition should be replaced.

Chipped Anvils

Sometimes an anvil can be chipped. Lower-energy specimens are affected the least amount because they are the hardest specimens and therefore have a more brittle fracture. The ductile high-energy specimens will produce higher than normal energy results, and the very ductile

super-high-energy specimens are affected most by a chipped anvil. This condition should be detected easily by a visual inspection before using the machine. New anvils are required when an anvil is chipped.

Anvil Relief

Some Charpy machine manufacturers have designed a machined relief at the bottom of the anvil. This anvil design does not meet the direct verification requirements of ASTM Standard E 23. The relief increases the measured energy for ductile high- and very-high-energy specimens. It can also cause twisting of the specimens, during fracture, that may also contribute to energy values higher than normal at all energy levels. However, this feature does not appear to add an excessive amount of energy to the test. (The results are usually within the allowed tolerances.)

Damaged Anvils

Under some test conditions, usually for elevated-temperature testing, the anvils can wear to a rough finish that creates excessive friction. This damaged condition is detected best on higher energy specimens. Damaged anvils usually cause the gouge marks to become wider and push the specimen material to form a ridge that can easily be detected with the fingernail. This damage usually causes artificially high high-energy results. Damaged anvils must be replaced.

Bent Pendulum

A pendulum bent in the direction of the swing produces gouge marks on a specimen. This gouge mark is usually deeper on the top edge of the specimen as it sits in the machine. The striker contacts the top edge of the specimen first, causing excessive tumbling and twisting. This excessive activity can cause the specimen to interact with the striker or the pendulum after fracture and create additional energy loss. A bent pendulum can be detected by placing an unbroken reference specimen in the machine and placing a piece of carbon paper on the surface opposite the notch. At this point, lightly tap the striker against the specimen. This will make a mark on the specimen that can be inspected. If the pendulum is not bent, the mark should appear the same width across the specimen. If the pendulum is bent, the mark will be wider at one edge and become thinner or even not visible at the other edge. We recommend that a new pendulum be installed on a machine with this problem.

Summary of Indirect Verification

Some aspects of Charpy machine condition and accuracy can be assessed only through the use of reference specimens. Further, some machine problems cause artificially low results while other machine problems cause artificially high results. In addition, deviations in procedures can cause similar results. These machine problems and procedural deviations may go undetected for years without some sort of physical check. For this reason, examination of the broken specimens is a critical part of the verification process. Many machine problems can be avoided or corrected by the use of the information presented in this paper. Also, suggested changes in procedure can help

to ensure a successful test. Verification specimens are available from various organizations around the world, including:

- the Institute for Reference Machines and Measurements (IRMM, in Belgium),
- Laboratoire National D'Essais (LNE, in France),
- National Institute for Standards and Technology (NIST, in the United States), and
- National Research Laboratory for Metals (NRLM, in Japan).

Summary

1. Recent refinements in the procedures continue to improve the accuracy of the test. Topical areas include the striker, anvils, specimens, and temperatures.

2. The Charpy scales used by the various NMIs are consistent, and the current round-robin promises further harmonization of the various procedures.

3. The history of past international interactions shows that a free and open interchange of ideas between countries is of benefit to all.

4. Direct and indirect verification testing is needed to assure the validity of data developed on a Charpy impact machine.

Acknowledgements

I appreciate the contributions of Dan Vigliotti and Chris McCowan at NIST to the procedures used to evaluate verification specimens.

References

1. Charpy, G. "Report on Impact Tests and the Work of Committee 26," (1912) In: Proceedings of the Sixth Congress of the International Association for Testing Materials, New York, IV, p. 1- 8, and the discussion of the report in Section A, p. 6-11.

2. Fink, D.A. (1989) In: Charpy Impact Test: Factors and Variables, ASTM STP 1072, J.M. Holt (Ed.), American Society for Testing and Materials, Philadelphia, Pennsylvania, p. 94-119.

3. Naniwa, T., Shibaike, M., Tanaka, M., Tani, H., Shiota, K., Hanawa, N., and Shiraishi, T. (1989) In: Charpy Impact Test: Factors and Variables, ASTM STP 1072, J.M. Holt (Ed.), American Society for Testing and Materials, Philadelphia, Pennsylvania, p. 67-80.

4. Nanstad, R.K., and Sokolov, M.A. (1995) In: Pendulum Impact Machines: Procedures and Specimens for Verification, ASTM STP 1248, T.A. Siewert and A.K. Schmieder, (Eds.), American Society for Testing and Materials, Philadelphia, p. 111-139.

5. Siewert, T. A., and Vigliotti, D. P. (1995) In: Pendulum Impact Machines: Procedures and Specimens for Verification, ASTM STP 1248, T.A. Siewert and A.K. Schmieder, (Eds.), American Society for Testing and Materials, Philadelphia, p. 140-152.

6. Ruth, E. A. (1995) In: Pendulum Impact Machines: Procedures and Specimens for Verification, ASTM STP 1248, T.A. Siewert and A.K. Schmieder, (Eds.), American Society for Testing and Materials, Philadelphia, p. 101-110.

7. Koester, R. D., and Barcus, S. E. (1989) In: Charpy Impact Test: Factors and Variables, ASTM STP 1072, J.M. Holt (Ed.), American Society for Testing and Materials, Philadelphia, Pennsylvania, p. 83-93.

8. Schmieder, A. K. (1989) In: Charpy Impact Test: Factors and Variables, ASTM STP 1072, J.M. Holt (Ed.), American Society for Testing and Materials, Philadelphia, Pennsylvania, p. 20-34.

9. Schmieder, A. K., Purtscher, P. T., and Vigliotti, D. P. (1995) In: Pendulum Impact Machines: Procedures and Specimens for Verification, ASTM STP 1248, T. A. Siewert and A. K. Schmieder, (Eds.), American Society for Testing and Materials, Philadelphia, Pennsylvania, p. 3-18.

10. Porro, F., Trippodo, R., Bertozzi, R., and Garagnani, R. "Impact Tester Compliance: Significance, Sensitivity, and Evaluation" (1989) In: Charpy Impact Test: Factors and Variables, ASTM STP 1072, J.M. Holt (Ed.), American Society for Testing and Materials, Philadelphia, Pennsylvania, p. 1-20.

11. Ruth, E. A., Vigliotti, D. P., and Siewert, T. A. (1995) In: Pendulum Impact Machines: Procedures and Specimens for Verification, ASTM STP 1248, T. A. Siewert and A. K. Schmieder, (Eds.), American Society for Testing and Materials, Philadelphia, Pennsylvania, p. 91-100.

12. Yamaguchi, Y., Takagi, S., and Nakano, H. (2000). In: Pendulum Impact Machines: A Century of Progress, ASTM STP 1380, T. A. Siewert and M. P. Manahan, Sr. (Eds.), American Society for Testing and Materials, Philadelphia, Pennsylvania, p. 164-180.

13. Alexander, D. J., and Klueh, R. L. (1989) In: Charpy Impact Test: Factors and Variables, ASTM STP 1072, J. M. Holt, (Ed). American Society for Testing and Materials, Philadelphia, Pennsylvania, p. 3-18.

14. Alexander, D. J., Corwin, W. R. and Owings, T. D. (1995) In: Pendulum Impact Machines: Procedures and Specimens for Verification, ASTM STP 1248, T. A. Siewert and A. K. Schmieder, (Ed). American Society for Testing and Materials, Philadelphia, Pennsylvania, p. 32-38.

15. Manahan, M. P., Stonesifer, R. B., Soong, Y., and Burger, J. M. (1995) In: Pendulum Impact Machines: Procedures and Specimens for Verification, ASTM STP 1248, T. A. Siewert and A. K. Schmieder, (Eds). American Society for Testing and Materials, Philadelphia, Pennsylvania, p. 39-69.

16. Marsh, F. J. (1995) In: Pendulum Impact Machines: Procedures and Specimens for Verification, ASTM STP 1248, T. A. Siewert and A. K. Schmieder, (Eds.), American Society for Testing and Materials, Philadelphia, Pennsylvania, p. 19-30.

17. Nanstad, R. K., Swain, R. L., and Berggren, R. G. (1989) In: Charpy Impact Test: Factors and Variables, ASTM STP 1072, J. M. Holt (Ed.), American Society for Testing and Materials, Philadelphia, Pennsylvania, p. 195-210.

18. Tobler, R. L., Reed, R. P., Hwang, I. S., Morra, M. M., Ballinger, R. G., Nakajima, H., and Shimamoto, S. (1991) In: J. Test. Eval. 19, 1, p. 34-40.

19. Manahan, Sr., M. P. (2000) In: Pendulum Impact Machines: A Century of Progress, ASTM STP 1380, T. A. Siewert and M. P. Manahan, Sr. (Eds.), American Society for Testing and Materials, Philadelphia, Pennsylvania, p. 286-297.

20. Sundqvist, M., and Chai, G. (2000) In: Pendulum Impact Machines: A Century of Progress, ASTM STP 1380, T. A. Siewert and M. P. Manahan, Sr. (Eds.), American Society for Testing and Materials, Philadelphia, Pennsylvania, p. 100-108.

21. Galban, G., Revise, G., Mougin, D., LaPorte, S. and LeFrancois, S. (2000) In: Pendulum Impact Machines: A Century of Progress, ASTM STP 1380, T.A. Siewert and M.P. Manahan, Sr. (Eds). American Society for Testing and Materials, Philadelphia, Pennsylvania, p. 109-133.

22. Driscoll, D.E. (1955) In: Symposium on Impact Testing, ASTM Special Publication 176, pp. 70-83.

23. McCowan, C.N., Pauwels, J., Revise, G., and Nakano, H. (2000) In: Pendulum Impact Testing: A Century of Progress, ASTM STP 1380, T.A. Siewert and M.P. Manahan, Sr., (Eds). American Society for Testing and Materials, West Conshohocken, Pennsylvania, p. 73-89.

24. McCowan, C.N. (2000) Harmonization of Measurement for Charpy Reference Materials. National Institute of Standards and Technology, Boulder, Colorado. p. 1-4.

25. Vigliotti, D.P., Siewert, T.A., and McCowan, C.N. (2000) Recommended Practice: Installing, Maintaining, and Verifying Your Charpy Impact Machine, NIST Technical Note 1500-8, National Institute of Standards and Technology, Boulder, Colorado. p. 1-17.

Part 2: Materials and Heat Treatment

C.N. McCowan,[1] D.P. Vigliotti,[1] and T.A. Siewert[1]

Evaluation of ABS Plastic Impact Verification Specimens

Reference: McCowan, C.N., Vigliotti, D.P., and Siewert, T.A., **"Evaluation of ABS Plastic Impact Verification Specimens,"** *Pendulum Impact Testing: A Century of Progress*, *ASTM STP 1380*, T. A. Siewert and M. P. Manahan, Sr., Eds., American Society for Testing and Materials, West Conshohocken, PA, 2000.

Abstract: Valid comparison of impact test energies reported by various laboratories over time depends on consistent performance of impact test machines. To learn what information verification specimens might reveal about the impact machines used to test plastics, and to learn the critical parameters that must be controlled to minimize the scatter in the data, we tested a batch of acrylonitrile-butadiene-styrene plastic (ABS plastic) specimens. Both Charpy and Izod sample configurations were tested. We concentrated on determining the variability in impact strength, but also studied the effects of room temperature aging, exposure to sunlight, and clamping pressure (for Izod testing) on the impact strength of the ABS plastic specimens. We conclude that: (1) the variability in impact strength for the ABS plastic impact specimens is comparable to that of metal specimens which are now used as verification specimens to test large-capacity impact machines; (2) the ABS plastic specimens did not show significant signs of aging during our tests; and (3) clamping pressures will have to be specified for verification testing of ABS plastic Izod specimens.

Keywords: Charpy impact testing, impact testing, Izod impact testing, plastic impact specimens, plastic impact verification specimens

[1] Materials Research Engineer, Engineering Technician, and Supervisory Metallurgist, respectively. NIST, materials Reliablility Division, 325 Broadway, Boulder, CO 80303.

Introduction

To help ensure valid comparisons of impact data, a requirement for the use of verification specimens was added to the Standard Test Methods for Notched Bar Impact Testing of Metallic Materials (ASTM E 23). This occurred because the metals impact testing community discovered that verification tests of impact machine performance were able to detect certain energy loss mechanisms that could not otherwise be observed during traditional physics-based measurements of machine performance (pendulum period, mass, mechanical friction, windage, etc.), and these mechanisms resulted in significant differences between the results of impact machines. This study evaluates the applicability of ABS plastic verification specimens (acrylonitrile-butadiene-styrene plastic) for testing the low-capacity impact machines used by the plastics industry. Nationally standardized verification specimens are not available for testing the performance of this class of impact machine.

Few studies on performance issues for plastics impact machines could be found. For our purposes, the most useful report was the one used to support the precision statement in ASTM (D 256 - 93a), "Standard Test Methods for Determining the Pendulum Impact Resistance of Notched Specimens." This report describes a round robin that included 6 different plastics and 25 laboratories.[1] It indicates that both the materials and the laboratories contribute significantly to the uncertainty in the data. This report also indicates that the impact strength for specimens produced from ABS plastic has low variability, as shown in **Table 1**.

The low coefficient of variation (CV) for the ABS plastic, which is the ratio of the standard deviation to the impact strength, compares well with the values of 0.04 typical of metal impact

verification specimens now being produced. So, in terms of variability, the ABS plastic appears to be a good candidate for use as an impact-verification material.

This report presents Izod and Charpy impact data for ABS plastic specimens to further confirm the low variability in impact strength. In addition, the effects of room-temperature aging, exposure to sunlight, and clamping pressure (for Izod testing) on the impact strength of the ABS plastic specimens are evaluated.

Table 1. Izod impact data summary from a previous study.[1] Only the laboratories that tested 10 specimens are included here.

Lab	Impact Strength		Standard Deviation		Standard Error		Coefficient of Variation
	J/m	(ft-lbf/in)	J/m	(ft-lbf/in)	J/m	(ft-lbf/in)	
1	599.97	11.24	24.02	0.45	7.47	0.14	0.04
2	557.80	10.45	9.07	0.17	2.67	0.05	0.02
4	591.96	11.09	10.68	0.20	3.20	0.06	0.02
5	580.22	10.87	13.88	0.26	4.27	0.08	0.02
6	564.74	10.58	13.35	0.25	4.27	0.08	0.02
7	639.47	11.98	17.08	0.32	5.34	0.10	0.03
9	583.96	10.94	13.35	0.25	4.27	0.08	0.02
10	587.69	11.01	28.29	0.53	9.07	0.17	0.05
11	614.38	11.51	14.41	0.27	4.80	0.09	0.02
12	584.49	10.95	19.75	0.37	5.34	0.12	0.03
13	547.13	10.25	20.82	0.39	5.34	0.12	0.04
14	545.52	10.22	19.22	0.36	5.87	0.11	0.04
23	541.25	10.14	8.01	0.15	2.67	0.05	0.01
24	547.12	10.25	8.54	0.16	2.67	0.05	0.02
Avg	577.55	10.82	15.48	0.29	4.80	0.09	0.03

Material and Procedures

All the Izod and Charpy samples used for this study were produced from the same batch of ABS plastic. They were cast in molds to produce geometries that meet the requirements of ASTM D 256. The notches were ground, and the dimensions of the specimens were measured prior to testing. In the case of Izod specimens, two specimens were cast at a time. These two specimens are distinguished from one another as "gate" and "dam" specimens, which refers to the end of the mold from which the specimens were taken.

Izod and Charpy specimens were made in thicknesses of 3.2 and 6.4 mm (0.125 and 0.25 in.) for the study. The measured thicknesses of the individual specimens were used to calculate the impact strength. All the notch depths and radii of the samples were found to be within the specifications of ASTM D 256. The 3.2 and 6.4 mm Izod samples were clamped in the specimen vise using 5 and 8.5 N·m (45 and 75 lbf-in.) of force respectively, applied using a torque wrench on a 12.7 mm diameter bolt (0.5 in. - 14).

All testing was performed at room temperature. The specimens were stored and tested at 21 ±1 °C (70 °F). The specimens were not conditioned at 50 ±5 % relative humidity prior to testing. The nominal humidity of the laboratory in which the specimens were stored and tested was low (less than 30 % relative humidity).

We used a pendulum impact machine with a maximum capacity of about 20 J. The machine was designed according to the requirements of D 256.

Results

As-Received Izod Specimens - The impact data for the as-received Izod specimens are given in **Table 2**. For the 6.4 mm (0.25 in.) thick specimens, the samples from the gate end of the mold had higher impact strength than the specimens from the dam end of the mold. The same trend was found for the 3.2 mm (0.125 in.) samples: the gate samples had higher impact strength than the dam samples. The differences due to specimen location were anticipated, and the results show that considering dam and gate specimens as separate groups helps to reduce unnecessary scatter for the material. The standard deviations in impact strength for the four groups of as-received specimens are low, confirming that the ABS plastic is a good candidate for use as an impact-verification material. The gate specimens had higher variation in impact strength than the dam specimens, but all the as-received specimens have CV values (0.04 or less) comparable to those of the steel specimens used for impact verification specimens.

As-Received Charpy Specimens - The tests results on the as-received Charpy specimens, given in **Table 3**, again show low variation in the impact strength for the ABS specimens. The 3.2 mm specimens have more variation in impact strength than the 6.4 mm specimens, but both thicknesses have fairly low CV values. A practical consideration here is that the thicker specimens may seat more firmly against the anvils in the Charpy test, and thus reduce the scatter in the test.

Table 2. *Initial Izod data for as-received impact specimens.*

ID	Cases	Mean Impact Strength J/m (ft-lbf/in)	Standard Deviation J/m (ft-lbf/in)	Coefficient of Variation
6.4 mm gate	25	222.48 (4.17)	7.91 (0.15)	0.04 (0.04)
6.4 mm dam	25	216.78 (4.06)	4.43 (0.08)	0.02 (0.02)
3.2 mm gate	25	233.21 (4.37)	6.83 (0.13)	0.03 (0.03)
3.2 mm dam	24	190.70 (3.57)	4.44 (0.08)	0.02 (0.02)

Table 3. *Data for the as-received Charpy impact specimens.*

ID	Cases	Mean Impact Strength J/m (ft-lbf/in)	Standard Deviation J/m (ft-lbf/in)	Coefficient of Variation
3.2 mm Charpy	25	241.54 (4.53)	12.26 (0.23)	0.05 (0.05)
6.4 mm Charpy	25	229.40 (4.30)	4.65 (0.09)	0.02 (0.02)

Room-Temperature Aging Data - The Izod data from the 30-, 90-, and 600-day aging groups and data from the as-received 3.2 mm gate specimens are given in **Table 4**. There appears to be a slight decrease in the impact strength of the specimens aged for 30 and 90 days, compared to the as-received specimens. However, the standard deviations of the three groups overlap, and the 30- and 90-day results are almost identical. The statistics for the three groups combined (0-, 30-, and 90-day) show only slightly more variation in impact strength than each group considered separately: (1) average impact strength of 228.5 J/m, (2) standard deviation of 7.37 J/m, and (3) coefficient of variation of 0.03.

Table 4. *Izod impact data for aged specimens. The as-received specimens are estimated to be 1 week old, 0 aging.*

ID	Aging (Days)	Cases	Mean Impact Strength J/m (ft-lbf/in)	Standard Deviation J/m (ft-lbf/in)	Coefficient of Variation
3.2 mm gate	0	25	233.21 (4.37)	6.83 (0.13)	0.03 (0.03)
3.2 mm gate	30	25	225.67 (4.23)	6.27 (0.12)	0.03 (0.03)
3.2 mm gate	90	24	226.60 (4.25)	6.48 (0.12)	0.03 (0.03)
3.2 mm gate	600	25	256.80 (4.81)	6.60 (0.12)	0.03 (0.03)

The impact strength of the specimens aged for 600 days increased by more than 10 % (delta %) and this indicates that aging had a significant effect on the impact strength of these ABS specimens. This result is somewhat surprising. Aging would generally be expected to degrade the impact properties of plastics. Unfortunately, since only one machine was used for these tests and the performance of this machine could not be checked with verification specimens, we are not convinced that the results of the 600-day tests are necessarily reliable enough to accept. These tests will have to be repeated in a more comprehensive study (with humidity control) to convincingly show the effects of aging on the impact strength of the ABS specimens, if any, and to better document any change in impact properties over time.

Specimens exposed to Sunlight - Izod and Charpy specimens were exposed to sunlight on a window sill and periodically rotated for 30 days prior to testing. The impact results, **Table 5**, show the average impact strength of the sunlight-exposed specimens is slightly lower than that of the as-received specimens. Again, the difference is not significant considering the standard deviations.

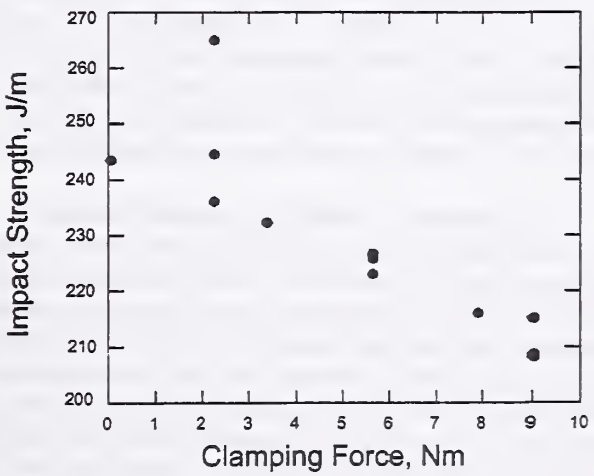
Table 5. *Izod Charpy impact data for specimens exposed to sunlight.*

ID	Exposure Time (Days)	Case	Mean Impact Strength J/m (ft-lbf/in)	Standard Deviation J/m (ft-lbf/in)	Coefficient of Variation
Charpy (3.2 mm)	0	25	241.54 (4.53)	12.26 (0.23)	0.05 (0.05)
Charpy (3.2 mm)	30	25	238.44 (4.47)	8.54 (0.16)	0.04 (0.04)
Izod (3.2 mm)	0	24	190.70 (3.57)	4.44 (0.08)	0.02 (0.02)
Izod (3.2 mm)	30	24	189.90 (3.56)	5.62 (0.11)	0.03 (0.03)

Table 6. *Izod clamping force.*

ID		Clamping Force N·m (in-lbs)	Mean Impact Strength J/m (ft-lbf/in)	Standard Deviation
3.2 mm gate	1	Just touching	243.38 (4.56)	
3.2 mm gate	1	1.1 (10)	217.09 (4.07)	
3.2 mm gate	5	2.3 (20)	247.69 (4.64)	10.77 (0.20)
3.2 mm gate	1	3.4 (30)	232.16 (4.35)	
3.2 mm gate	1	4.5 (40)	235.06 (4.40)	
3.2 mm gate	4	5.7 (50)	225.44 (4.22)	1.66 (0.03)
3.2 mm gate	1	6.8 (60)	216.96 (4.07)	
3.2 mm gate	1	7.9 (70)	215.91 (4.05)	
3.2 mm gate	3	9.0 (80)	210.75 (3.95)	3.74 (0.07)

Clamping Force in Izod Tests - The force with which the ABS specimens were clamped for Izod testing had a significant effect on the impact strength of the specimens, as shown in **Figure 1** and **Table 6**. At low clamping forces (less than 3 N·m), the scatter in the results and the impact strength increased. At clamping forces greater than 6 N·m, it appears that the effect of clamping on scatter in the impact strength decreases.



Discussion

The average impact strengths for the groups of ABS plastic impact specimens that were tested ranged between about 190 and 240 J/m. Converting to impact energy, the range is from about 0.6 to 0.8 J. This is a useful energy for a low-energy verification specimen, considering that the machines used to test plastics have low capacities. The impact machine used for this study, for example, has a low range of about 0 to 3 J, and weights can be added to the pendulum to increase the capacity up to about 20 J.

The variation in impact strength for the ABS specimens was low, and this is important for a material being considered for use as a verification specimen. The variation must be low enough to allow a small sample size to be used when testing the performance of a machine. Typically the CV values for the ABS material were 0.04 or less, and previous experience with metal impact verification specimens indicates that this variability would allow a sample size of around 5. The sample size for a verification test depends on the specific requirements of the test, but clearly the ABS plastic has low variation.

The age of the ABS plastic specimens did not have a significant effect on the impact strength for short aging times, but the impact strength increased for the specimens that were aged for 600 days. These results imply that the impact strength of the specimens does not remain stable over time, but further testing is required to confirm these results.

The impact strength of the specimens did not appear to be too sensitive to sunlight exposure. The sunlight exposure tests were somewhat simplistic, but they indicate that the specimens could be left out on a sunlit desk for several days prior to testing without adverse effect.

The impact strength of the ABS Izod specimens decreased significantly as clamping pressure was increased on the specimen. This is apparently a common effect for some plastics, so we included these preliminary tests here to help determine how much the clamping pressure affected the impact strength of the ABS plastic specimens (D 256, note 7). The results show that clamping pressure is an important variable and would have to be specified for ABS Izod verification specimens.

Overall, the impact strength of the ABS specimens was consistent and this makes the material a good candidate for the production of verification specimens for low capacity impact machines. The affect of aging on the ABS plastic will have to be considered and there are other considerations in choosing a specimen to verify the performance of impact machines. A verification specimen should test specific aspects of the machine performance that are difficult or impossible to verify by static measurement alone.

Experience in verifying the performance of large capacity impact machines, used for metals, has shown that testing hard specimens (HRC 45 or greater) is the best way to verify that a machine is properly mounted. The very rapid loading of the machine associated with the impact of these hard specimens is needed to reveal a mounting problem for the machine. Softer specimens, on the other hand, are found to better show the effects of anvil and striker condition on the machine performance. For the lower-capacity impact machines used to test plastics, factors that affect the test results and the machine material interactions should be carefully considered to help ensure that appropriate and useful verification specimens are chosen. It is likely that several different types of specimens will be needed to adequately verify the performance of the low-capacity impact machines.

Conclusions

(1) The variability in impact strength for the ABS plastic impact specimens is comparable to that of metal specimens that are presently used as verification specimens to test large-capacity impact machines.

(2) Clamping pressures will have to be specified for verification testing of ABS plastic Izod specimens.

(3) The effects of aging and relative humidity need more study.

Acknowledgments

We are grateful to Dave Scarlett of Bayer Corporation for producing the ABS impact specimens used in this study.

References

1. Research Report RR: D20-1134, available from ASTM Headquarters, Conshohocken, PA.

Appendix I

The Groups are identified as follows:
- *I* indicates initial, as-received, test group;
- *a* indicates aged for either 30 or 90 days (a30 or a90);
- *4* and *8* indicate specimen thickness (4 = 6.4 mm or 1/4 in. and 8 = 3.2 mm or 1/8 in.);
- *uv* indicates exposure to sunlight.

Table A1.1- *Izod Data*

Case	Group	ID	J/m	ft-lbf/in
1	a30g8	g101	220.238	4.126
2	a30g8	g102	228.511	4.281
3	a30g8	g103	223.253	4.182
4	a30g8	g104	217.515	4.075
5	a30g8	g105	225.469	4.224
6	a30g8	g106	219.864	4.119
7	a30g8	g107	230.059	4.310
8	a30g8	g108	228.511	4.281
9	a30g8	g109	228.031	4.272
10	a30g8	g110	222.586	4.170
11	a30g8	g111	223.440	4.186
12	a30g8	g112	226.002	4.234
13	a30g8	g113	225.469	4.224
14	a30g8	g114	215.754	4.042
15	a30g8	g115	235.984	4.421
16	a30g8	g116	229.952	4.308
17	a30g8	g117	232.354	4.353
18	a30g8	g118	222.907	4.176
19	a30g8	g119	241.322	4.521
20	a30g8	g120	215.807	4.043
21	a30g8	g121	224.775	4.211
22	a30g8	g122	221.732	4.154
23	a30g8	g123	219.330	4.109
24	a30g8	g124	228.938	4.289
25	a30g8	g125	233.849	4.381
26	a90g8	g126	226.590	4.245
27	a90g8	g127	221.305	4.146
28	a90g8	g128	225.682	4.228
29	a90g8	g129	234.169	4.387
30	a90g8	g130	232.515	4.356
31	a90g8	g131	220.131	4.124
32	a90g8	g132	232.301	4.352
33	a90g8	g133	230.913	4.326
34	a90g8	g134	224.988	4.215
35	a90g8	g135	220.932	4.139
36	a90g8	g136	222.319	4.165
37	a90g8	g137	221.625	4.152
38	a90g8	g138	230.273	4.314
39	a90g8	g139	241.856	4.531
40	a90g8	g140	215.594	4.039
41	a90g8	g141	231.767	4.342
42	a90g8	g142	230.593	4.320
43	a90g8	g143	229.579	4.301
44	a90g8	g144	230.913	4.326
45	a90	g145	220.771	4.136
46	a90g8	g147	229.045	4.291
47	a90g8	g148	214.366	4.016

Case	Group	ID	J/m	ft-lbf/in
48	a90g8	g149	220.451	4.130
49	a90g8	g150	229.792	4.305
50	id4	d1	210.630	3.946
51	id4	d10	215.647	4.040
52	id4	d11	214.206	4.013
53	id4	d12	213.459	3.999
54	id4	d13	213.939	4.008
55	id4	d14	220.558	4.132
56	id4	d15	216.181	4.050
57	id4	d16	212.925	3.989
58	id4	d17	219.811	4.118
59	id4	d18	216.448	4.055
60	id4	d19	225.629	4.227
61	id4	d2	214.153	4.012
62	id4	d20	218.636	4.096
63	id4	d21	219.010	4.103
64	id4	d22	226.323	4.240
65	id4	d23	223.974	4.196
66	id4	d24	219.811	4.118
67	id4	d25	212.605	3.983
68	id4	d3	210.896	3.951
69	id4	d4	218.209	4.088
70	id4	d5	216.448	4.055
71	id4	d6	220.398	4.129
72	id4	d7	214.206	4.013
73	id4	d8	215.220	4.032
74	id4	d9	210.096	3.936
75	id8	d26	193.068	3.617
76	id8	d27	197.605	3.702
77	id8	d28	183.460	3.437
78	id8	d29	188.424	3.530
79	id8	d30	194.776	3.649
80	id8	d31	193.762	3.630
81	id8	d32	183.567	3.439
82	id8	d33	185.435	3.474
83	id8	d34	188.584	3.533
84	id8	d35	.	.
85	id8	d36	187.090	3.505
86	id8	d37	198.940	3.727
87	id8	d38	196.751	3.686
88	id8	d39	187.677	3.516
89	id8	d40	188.157	3.525
90	id8	d41	188.958	3.540
91	id8	d42	185.489	3.475
92	id8	d43	189.972	3.559
93	id8	d44	191.307	3.584
94	id8	d45	191.307	3.584
95	id8	d46	191.894	3.595

Case	Group	ID	J/m	ft-lbf/in
96	id8	d47	190.613	3.571
97	id8	d48	197.605	3.702
98	id8	d49	188.424	3.530
99	id8	d50	194.563	3.645
100	ig4	g1	226.750	4.248
101	ig4	g10	215.647	4.040
102	ig4	g11	214.046	4.010
103	ig4	g12	213.512	4.000
104	ig4	g13	214.046	4.010
105	ig4	g14	231.554	4.338
106	ig4	g15	223.761	4.192
107	ig4	g16	226.109	4.236
108	ig4	g17	233.048	4.366
109	ig4	g18	227.871	4.269
110	ig4	g19	232.942	4.364
111	ig4	g2	231.821	4.343
112	ig4	g20	225.415	4.223
113	ig4	g21	231.554	4.338
114	ig4	g22	232.301	4.352
115	ig4	g23	231.340	4.334
116	ig4	g24	227.390	4.260
117	ig4	g25	217.729	4.079
118	ig4	g3	210.843	3.950
119	ig4	g4	218.316	4.090
120	ig4	g5	216.715	4.06[illegible]0
121	ig4	g6	219.917	4.12[illegible]0
122	ig4	g7	214.046	4.01 0
123	ig4	g8	215.113	4.03[illegible]0
124	ig4	g9	210.309	3.94[illegible]0
125	ig8	g26	226.910	4.25[illegible]1
126	ig8	g27	231.821	4.34[illegible]3
127	ig8	g28	243.350	4.55[illegible]9
128	ig8	g29	243.831	4.56[illegible]8
129	ig8	g31	236.625	4.43[illegible]3
130	ig8	g32	229.632	4.30[illegible]2
131	ig8	g33	245.058	4.59[illegible]1
132	ig8	g34	228.618	4.28[illegible]3
133	ig8	g35	230.326	4.31[illegible]5
134	ig8	g36	234.308	4.39[illegible]0
135	ig8	g37	227.908	4.27[illegible]0
136	ig8	g38	228.116	4.27[illegible]4
137	ig8	g39	243.649	4.56[illegible]5
138	ig8	g40	235.002	4.40[illegible]3
139	ig8	g41	241.744	4.52[illegible]9
140	ig8	g42	230.524	4.31[illegible]9
141	ig8	g43	225.373	4.22[illegible]2
142	ig8	g44	227.086	4.25[illegible]4
143	ig8	g45	228.591	4.28[illegible]2

Case	Group	ID	J/m	ft-lbf/in
144	ig8	g46	221.780	4.155
145	ig8	g47	222.800	4.174
146	ig8	g48	234.484	4.393
147	ig8	g49	238.802	4.474
148	ig8	g50	237.762	4.454
149	ig8	g60	236.198	4.425
150	uvd8	d101	192.641	3.609
151	uvd8	d102	179.350	3.360
152	uvd8	d103	183.247	3.433
153	uvd8	d105	190.506	3.569
154	uvd8	d106	187.997	3.522
155	uvd8	d107	192.961	3.615
156	uvd8	d108	195.150	3.656
157	uvd8	d109	201.075	3.767
158	uvd8	d110	190.666	3.572
159	uvd8	d111	191.307	3.584
160	uvd8	d112	193.655	3.628
161	uvd8	d113	191.627	3.590
162	uvd8	d114	185.542	3.476
163	uvd8	d115	194.616	3.646
164	uvd8	d116	178.816	3.350
165	uvd8	d117	194.456	3.643
166	uvd8	d118	192.001	3.597
167	uvd8	d119	182.766	3.424
168	uvd8	d120	185.382	3.473
169	uvd8	d121	189.812	3.556
170	uvd8	d122	199.943	3.746
171	uvd8	d123	186.770	3.499
172	uvd8	d124	190.026	3.560
173	uvd8	d125	187.357	3.510

Table A1.2: Charpy Data

Case	Group	ID	J/m	ft-lbf/in
1	i4	76.000	227.123	4.255
2	i4	77.000	227.817	4.268
3	i4	78.000	223.280	4.183
4	i4	79.000	225.522	4.225
5	i4	80.000	226.483	4.243
6	i4	81.000	227.550	4.263
7	i4	82.000	224.134	4.199
8	i4	83.000	226.643	4.246
9	i4	84.000	225.415	4.223
10	i4	85.000	222.373	4.166
11	i4	86.000	241.322	4.521
12	i4	87.000	226.109	4.236
13	i4	88.000	232.568	4.357
14	i4	89.000	236.358	4.428
15	i4	90.000	234.970	4.402
16	i4	91.000	229.152	4.293
17	i4	92.000	227.924	4.270
18	i4	93.000	229.098	4.292
19	i4	94.000	229.899	4.307
20	i4	95.000	237.425	4.448
21	i4	96.000	232.568	4.357
22	i4	97.000	227.390	4.260
23	i4	98.000	230.326	4.315
24	i4	99.000	234.169	4.387
25	i4	100.000	229.365	4.297
26	i8	51.000	219.864	4.119
27	i8	52.000	239.347	4.484
28	i8	53.000	250.556	4.694
29	i8	54.000	234.276	4.389
30	i8	55.000	251.731	4.716
31	i8	56.000	248.475	4.655
32	i8	57.000	235.557	4.413
33	i8	58.000	254.880	4.775
34	i8	59.000	244.151	4.574
35	i8	60.000	207.907	3.895
36	i8	61.000	241.375	4.522
37	i8	62.000	233.155	4.368
38	i8	63.000	239.827	4.493
39	i8	64.000	245.165	4.593
40	i8	65.000	234.116	4.386
41	i8	66.000	224.134	4.199
42	i8	67.000	254.346	4.765
43	i8	68.000	238.813	4.474
44	i8	69.000	253.118	4.742
45	i8	70.000	258.136	4.836
46	i8	71.000	252.958	4.739
47	i8	72.000	229.205	4.294
48	i8	73.000	249.702	4.678

Case	Group	ID	J/m	ft-lbf/in
49	i8	74.000	251.090	4.704
50	i8	75.000	246.713	4.622
51	uv	251.000	231.447	4.336
52	uv	252.000	242.283	4.539
53	uv	253.000	241.429	4.523
54	uv	254.000	223.440	4.186
55	uv	255.000	241.535	4.525
56	uv	256.000	221.732	4.154
57	uv	257.000	243.777	4.567
58	uv	258.000	249.862	4.681
59	uv	259.000	252.104	4.723
60	uv	260.000	243.991	4.571
61	uv	261.000	241.749	4.529
62	uv	262.000	236.411	4.429
63	uv	263.000	237.585	4.451
64	uv	264.000	247.087	4.629
65	uv	265.000	246.019	4.609
66	uv	266.000	237.799	4.455
67	uv	267.000	239.721	4.491
68	uv	268.000	223.067	4.179
69	uv	269.000	238.706	4.472
70	uv	270.000	229.846	4.306
71	uv	271.000	237.372	4.447
72	uv	272.000	252.425	4.729
73	uv	273.000	238.706	4.472
74	uv	274.000	227.337	4.259
75	uv	275.000	235.450	4.411

T.A. Siewert, D.P. Vigliotti, L.B. Dirling, and C.N. McCowan

Performance Verification of Impact Machines for Testing Plastics

NIST Journal of Research, 104(6), Nov.-Dec. 1999, pp. 557-565.

Abstract: Valid comparison of impact test energies reported by various organizations and over time depends on consistent performance of impact test machines. This paper investigates the influence of various specimen and test parameters on impact energies in the 1 J to 2 J range for both Charpy V-notch and Izod procedures, leading toward the identification of a suitable material for use in a program to verify machine performance. We investigated the influences on the absorbed energy of machine design, test material, specimen cross-sectional area, and machine energy range. For comparison to published round-robin data on common plastics, this study used some common metallic alloys, including those used in the international verification program for metals impact machines and in informal calibration programs of tensile machines. The alloys that were evaluated include AISI type 4340 steel, and five aluminum alloys: 2014-T6, 2024-T351, 2219-T87, 6061-T6, and 7075-T6. We found that certain metallic alloys have coefficients of variation comparable to those of the best plastics that are reported in the literature. Also, we found that the differences in absorbed energy between two designs of machines are smaller than the differences that can be attributed to the specimens alone.

Key Words: aluminum; Charpy impact test; Izod impact test; plastics; steel; verification; V-notch

1. Introduction

About 40 years ago, a requirement for the use of verification specimens was added to the standard for impact testing of metals, American Society for Testing and Materials (ASTM) Standard E 23 [1]. This occurred because the metals impact testing community discovered that verification tests of impact-machine performance using reference specimens were able to detect certain energy-loss mechanisms, mechanisms that could not otherwise be observed during traditional physics-based measurements of machine performance (pendulum period, mass, mechanical friction, windage, etc). This present paper evaluates the use of verification specimens for machines used to test plastics, and suggests what information these specimens provide about machine performance.

Few studies on performance issues for plastics impact machines could be found. For our purposes, one of the most useful was the one used to support the precision statement in ASTM Standard D 256, "Standard Test Methods for Impact Resistance of Plastics and Electrical Insulating Materials" [2]. That report describes a round robin that included six different plastics and 25 different laboratories. It indicates that both the materials and the laboratories make significant contributions to the uncertainty in the data. Another study, also an ASTM research report, indicates that the effect of notch radius (for plastic materials) is linear over the range of notch radii of 0.03 mm to 2.5 mm [3].

2. Procedure

2.1. Material Selection

The first material to be included in the test plan was AISI type 4340 steel (of a special, high-purity grade) since this has been used for many years to make verification specimens for metals impact machines. Therefore, it serves as a good benchmark against which other materials can be measured. We compared this 4340 steel to several aluminum alloys: 2014-T6, 2024-T351, 2219-T87, 6061-T6, and 7075-T6. Alloy 6061-T6 was selected because of its reproducible performance in some informal tensile testing programs. The other aluminum alloys were selected because they are readily available, and also because they are known to possess a good blend of strength and ductility in structural applications. All these aluminum alloys have a lower modulus (stiffness), about 70 GPa, than that of the 4340 steel, about 200 GPa [4]. This means that they deform at a force lower than for steels, yet still at a force much greater than for plastics, whose moduli usually fall between 2 GPa and 12 GPa [5]. A higher modulus in the verification test material is not necessarily detrimental, since it can serve to better identify an energy loss mechanism in an impact machine that is due to internal friction in the components that are loaded during fracture. The larger oscillation during these higher loads with metal specimens helps us to determine whether this effect is significant during routine testing with plastic specimens.

Stability of impact energy over time is one of the most desirable features in the verification specimens used to assess machine performance. The 4340 steel has a shelf life of at least several years, and so is a good benchmark against which other materials can be measured. Although aluminum alloys such as the 2000, 6000, and 7000 series age harden, these effects were minimized through careful selection of alloy and lots. The 6061 and 7075 alloys were treated at elevated temperature (although lower than the tempering treatments for the 4340 steel), and the steep reduction in diffusion rate with lower temperature drastically limits subsequent aging at room temperature. The 2000 series alloys will age at room temperature, but 80 % to 90 % of the hardening is completed in 4 to 5 days [6]. To minimize the small amount of residual aging, we took our specimens from bars and plates that had been in inventory for several years. In summary, we expect very little change over time in the mechanical properties of specimens made from these metallic materials.

Another desirable feature for verification specimens is complete fracture of the specimens during impact. Complete fracture is preferred because we can compare the marks on both fractured halves. The specimens develop marks during the initial strike (when the pendulum hits the specimen), during fracture, and during subsequent collisions as the specimens leave the machines. Assessment of these marks during the post-test evaluation of the specimens (part of the metals impact test ASTM Standard E 23 procedures) provides guidance to machine owners about alignment problems and wear.

We did not include any plastics in this evaluation, because sufficient reference data exist in the two reports cited in the introduction. The goal of this study was to compare the data for these candidate metals to the existing data for the plastics.

2.2. Specimen Design and Preparation

We selected a specimen configuration (**Fig. 1**) that was designed to supplement the machine verification tests described in International Organization for Standardization (ISO) Draft International Standard (DIS) 13802 [7]. The dimensions were selected based on the plastic specimens described in ASTM Standard D256, but were changed where necessary to allow for the different materials properties of the metals. We permitted deviations from the standard configuration, such as side grooves on each side of the notch in some specimens, and allowing the length to be that specified either for Charpy or for Izod impact testing. One of the goals of this study was to determine the energies that would be developed by specimens of various sizes. The approach taken in the metals impact standard is to verify the machine performance, using specimens distributed over the useful range of the machine (up to 80 % of the machine capacity). For a 22 J impact machine, this means specimens with energies from 1 J or 2 J, up to about 15 J.

The standard configuration for the type 4340 steel specimens used in metal impact machines (as described in ASTM Standard E 23) has a cross section of 10 mm by 10 mm [1]. When heat-treated to produce a low energy, this specimen configuration absorbs about 15 J of energy from the pendulum. This 15 J specimen seemed appropriate to evaluate the performance of our plastics impact machine when configured for its maximum capacity of 22 J. However, our machine was damaged in an attempt to break one of these specimens in some preliminary tests. Apparently, our machine is able to tolerate energies in this range only when the fracture event is spread over a longer time, such as when the low modulus plastic specimens deform before fracture. The impact energy is a single number that is the integral of the incremental resistance of the specimen to fracture as the pendulum swings through its range. Therefore, by itself, the impact energy reading is an inaccurate way to compare the responses of low- and high-modulus materials, since it does not reflect the influence of the maximum load on the machine-specimen interactions.

After the impact machine was repaired, we continued our initial evaluations (to establish the experiment design) using miniature specimens designed to evaluate only the lower end of the machine range, between 1 J and 2 J. We selected steel specimen cross sections of 5 mm by 5 mm, 4 mm by 5 mm (notched across the 4 mm face), and 4 mm by 4 mm for some preliminary tests. This size range was designed to determine the cross section that would produce energies within the desired range, as well as a specimen size that would completely fracture upon impact. Another reason for concentrating on specimens of lower energy in the rest of this paper is that metal specimens show more ductility in thinner sections. Future tests with larger specimens, for the higher end of the machine capacity, should exhibit a more brittle fracture.

The impact data from the four specimens tested with each of the three cross sections, 4 mm by 4 mm, 4 mm by 5 mm, and 5 mm by 5 mm, were compared to the cross-sectional area. To obtain the cross-sectional area, we multiplied the two dimensions and subtracted the area of the notched region (1 mm deep) from the product. We found a linear relationship between cross-sectional area and energy (at least for this limited energy range), then used these data to standardize on a specimen dimension of 4 mm by 5 mm for the majority of our tests, since this size fell near the center of the desired energy range.

Unfortunately, the specimens in these preliminary tests did not completely separate into halves upon impact. Even the 5 mm by 5 mm specimens left the machine with the two halves still joined by a thin ligament, but bent by the impact to about a 90° angle. The ligament was so thin that the specimen halves could be bent to a 180° angle with two fingers, at which point the two halves would separate. Therefore, the specimen absorbed almost all the energy needed to fracture it (and also absorbed the kinetic (toss) energy imparted by the striker), but we did not gain data on possible jamming between the striker and anvils that might occur as broken halves left the machine. While the standard 10 mm by 10 mm specimen in a metals impact test shows almost no ductility and leaves the machine in two halves traveling at high velocity, all specimens in the range of 4 mm by 4 mm to 5 mm by 5 mm showed substantial ductility. Therefore, an optimal cross section is nearer to 10 mm by 10 mm, but we decided not to increase the section size since it would increase the required energy for fracture beyond the desired range.

A thorough evaluation of all variables calls for a full factorial experimental program with a large number of replicate tests, which was beyond the scope, budget, and time available for this study. Rather, this study is an initial evaluation to determine which variables might be most important and should provide the basis for selection of the variables to include in a future round robin. Nevertheless, we have used statistical summaries of the data using common formulas to give some estimates of repeatability.

2.3. Machine design

Our plastics impact machine can evaluate both Charpy V-notch and Izod specimens, which allowed us to develop data in the two different test configurations. The number of specimens tested with the Izod technique was much smaller than that tested with the Charpy technique, and was sufficient only to estimate whether there were some differences in the coefficients of variation for these specimens between the two configurations. Although many companies follow the Charpy V-notch and Izod impact test procedures described in ASTM Standard D 256, we have seen a growing interest in ISO Standard DIS 13802, and decided to follow its procedural requirements. Differences between these two standards include anvil spacing, included angle and radius of the striker tip, anvil radius, and many other details of the test procedure. These differences preclude direct comparison of the means between specimens tested according to the two standards, but we believe that the differences are sufficiently minor to conclude that the standard deviations and coefficients of variation (CVs, or relative standard deviations) should be about the same.

While most tests were performed using a single machine set up for a full-scale capacity of 22 J, we also removed the masses bolted to the pendulum, reducing the machine range for several tests to 5 J. We also performed some tests using a conventional metals impact machine with a capacity of 358 J, and using another design of plastics impact machine configured for a capacity of 22 J.

3. Results and Discussion

All of the data from our tests are included in **Table A-1** (Charpy) and **Table A-2** (Izod) in the appendix. This body of data consists of sets of similar specimens, typically sets of 3 to 10

specimens each, so we could evaluate the scatter between the specimens tested at the same set of conditions. These data are summarized in the appendix in **Table A-3**, which combines the data in each set.

Tables A-1 and A-2 contain columns that describe all of the parameters that were varied or measured in the evaluation. Even when summarized in **Table A-3**, this large body of data makes comparisons difficult, so we have selected subsets of the data in the following discussion to highlight the effects of the different parameters. These subsets exclude the columns with data that were held constant (listed in the notes below each table), but add columns with the calculated standard deviations and coefficients of variation. For the 4340 steel, comparisons are made only between data from sets that were heat treated in the same batch (with similar prefixes).

3.1. Test Material

Table 1 includes summary data for the six materials (one steel and five aluminum alloys) included in this investigation, and is intended primarily to compare the repeatabilities of the data for different materials. The statistical data should be used with care in comparing the different materials, since the values are based on only three specimens for each material. Also, standard deviations cannot be compared fairly when the means are different. Therefore, the next column lists the coefficients of variation (CVs), which are the standard deviations divided by the means, and which are also commonly called relative standard deviations. These permit easier comparison of repeatability among specimen types and shapes with different means, but they do also suffer from the same statistical deficiency that comes from having only three specimens. We can thus make only broad generalizations about the data.

The data for the metals seem to fit into three groups according to the coefficient of variation: CV up to 0.036, CV near 0.05, and CV near 0.1. We compared these to the interlaboratory data reported for plastics in Table 1 of ASTM Standard D256-93a [2]. These data fell into two groups: coefficients of variation between 0.042 and 0.058 for the plastics with lower energy absorption (phenolic, acetal, reinforced nylon, and polypropylene), and coefficients of variation between 0.012 and 0.018 for the plastics with higher energy absorption (ABS and polycarbonate). The summary of these data in ASTM Standard D256 does not mention the material thickness; instead, it uses the usual plastics convention of normalizing the energy to 25 mm of specimen width (notch length). This facilitates comparison of plastics of different sheet thicknesses (often 3 mm to 12.7 mm), but makes analysis of the data in terms of machine energy range more difficult. Nevertheless, for a given machine range, the plastics with lower energy absorption will obviously yield data that are at the lower end of that machine's range.

The ductile plastics have low CVs and so should be suitable for assessing machine repeatability at the high end of the machine range. At the low end of the machine range, the metals with the lower CVs (especially 4340 steel) beat the best of the plastics, by about a factor of two. Therefore, at the low end of the machine capacity, metals offer the possibility of at least matching the ability of specimens of plastics in resolving machine repeatability or the source of

machine uncertainties. In addition, metal specimens put a larger load on the machine striker and frame, better revealing mounting and other structural problems.

3.2. Side Grooves

Table 2 shows that side grooves reduce the mean energy (by an amount much greater than that explained by the reduction in cross-sectional area), but the coefficient of variation either stayed about the same or increased slightly. Also, the specimen halves were still joined by a similarly sized ligament. These unexpected results do not support more than a few confirming tests on aluminum alloys in any future evaluation.

3.3. Notch Radius

Table 3 shows the effect of halving the notch radius from 0.5 mm to 0.25 mm. In general, notches are stress concentrators that reduce the ability of a material to sustain a load, and so notches promote brittle, rather than ductile, failure [8]. In addition, sharper notches should decrease the scatter, as a sharper notch increases the local stress at the crack, and its variable contribution to the scatter. However, these data show that the consistency was much worse with the sharper notch, although the absorbed energy indeed decreased by about 10 %. Once again, these disappointing results seem to minimize the value of including a large number of tests for this variable in any future evaluations.

3.4. Machine Capacity and Design

Table 4 shows the effect of machine capacity and design. There was as much as 45 % variation in the energy as the machine capacity was changed from 22 J to 358 J, at least for alloy 7075. This variation is several times greater than the standard deviation and so appears to be significant. However, the effect does not seem significant for alloy 6061-T6, with the effect being less than the standard deviation. This lack of significance is also evident for changes in the machine range from 5 J to 22 J, comparison of plastics impact machines from two different manufacturers, and comparison between the plastics impact machines (with capacities of 22 J) and a metals impact machine (with a capacity of 358 J). In retrospect, it would have been better to have made more specimens of 4340 and repeated this test using specimens having a smaller standard deviation. However, both these results support the robustness of the pendulum impact-machine concept and indicate that it has been implemented consistently in these different machine designs.

3.5. Charpy versus Izod

The limited data prevent us from drawing strong conclusions about the effect of test orientation (Izod versus Charpy), but an informal comparison of the two types of data summarized in Table A-3 reveals no clear distinction between the two. It seems as though the difference in mean energy between Izod and Charpy tests is less than the variability due to other test parameters and so cannot be resolved. The same statement can be made for the CVs.

4. Conclusions

1. At low energies (low end of the machine capacity), certain metallic alloys have CVs that are about half that of the plastics in this range. This seems to support the option of using metal impact specimens to verify the performance of plastics impact machines at the low end of their range.

2. Metal specimens would increase the load on the machine striker and frame, permitting better resolution of problems with machine rigidity and mounting.

3. Verification testing provides valuable performance data. For example, the data for alloy 6061-T6 indicate that two different designs of impact machines for plastics produce results that agree within the uncertainty that can be attributed to the specimens alone.

5. Acknowledgment

We appreciate the assistance of Jennifer Caragol in revising and formatting the tables.

6. References

[1] ASTM Standard E 23, Standard Test Methods for Notched Bar Impact Testing of Metallic Materials, Vol 03.01, American Society for Testing and Materials, Conshohocken, Pennsylvania, (1999).

[2] Research Report RR: D20-1134, untitled, available from American Society for Testing and Materials, Conshohocken, Pennsylvania.

[3] Research Report RR: D20-1021, untitled, available from American Society for Testing and Materials, Conshohocken, Pennsylvania.

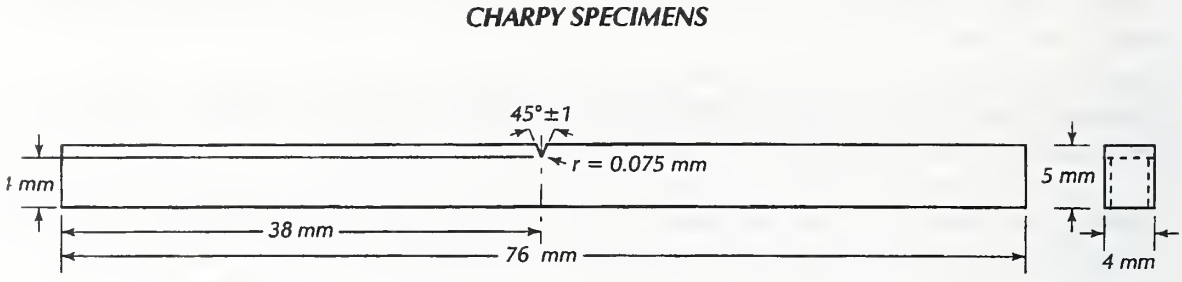
[4] Smithell's Metals Reference Book, Sixth Ed., E.A. Brandes, Butterworths, London, (1983) p. 15-3.

[5] Guide to Selecting Engineered Materials, Advanced Materials and Processing, Vol. 137 (6), ASM International, Materials Park, Ohio, June (1990) p. 23.

[6] Metals Handbook Ninth Edition, (Vol. 2) - Properties and Selection: Nonferrous Alloys and Pure Metals, ASM International, Materials Park, Ohio, (1979) p. 31.

[7] ISO Standard DIS 13802, Plastics - Pendulum impact-testing machines - Validation of Charpy, Izod, and tensile impact-testing machines, available from the ISO secretariat AFNOR, (1995).

[8] Metals Handbook Ninth Edition, (Vol. 11) - Failure Analysis and Prevention, ASM International, Materials Park, Ohio, (1986) p. 85.



NOTE: When used, sides were grooved with 45° cutter and 0.25 mm radius

Figure 1. Charpy specimen design.

Table A-1. Data From Charpy Tests.

SPECIMEN ID	MATERIAL	ABSORBED ENERGY (J)	NOTCH RADIUS (mm)	SPECIMEN DIMENSIONS (mm)	POTENTIAL ENERGY (J)	COMMENTS
20	6061-T6	3.576	0.50	4 by 5	5	Remaining Ligament
21	6061-T6	3.648	0.50	4 by 5	5	Remaining Ligament
22	6061-T6	3.338	0.50	4 by 5	5	Complete Breaks
23	6061-T6	3.301	0.50	4 by 5	22	Remaining Ligament
24	6061-T6	3.571	0.50	4 by 5	22	Remaining Ligament
25	6061-T6	3.550	0.50	4 by 5	22	Complete Breaks
26	6061-T6	3.036	0.25	4 by 5	22	Complete Breaks
27	6061-T6	3.435	0.25	4 by 5	22	Complete Breaks
28	6061-T6	2.858	0.25	4 by 5	22	Complete Breaks
29	7075-T6	1.138	0.25	4 by 5	22	Complete Laminate Breaks
30	7075-T6	1.138	0.25	4 by 5	22	Complete Laminate Breaks
31	7075-T6	1.238	0.25	4 by 5	22	Complete Laminate Breaks
32	2219-T87	1.245	0.25	4 by 5	22	Complete Brittle Fracture
33	2219-T87	1.320	0.25	4 by 5	22	Complete Brittle Fracture
34	2219-T87	1.382	0.25	4 by 5	22	Complete Brittle Fracture
35	2014-T6	2.889	0.25	4 by 5	22	Complete Laminated Fracture
36	2014-T6	3.000	0.25	4 by 5	22	Complete Laminated Fracture
37	2014-T6	2.791	0.25	4 by 5	22	Complete Laminated Fracture
38	2024-T351	1.970	0.25	4 by 5	22	Complete Brittle Fracture
39	2024-T351	1.828	0.25	4 by 5	22	Complete Brittle Fracture

SPECIMEN ID	MATERIAL	ABSORBED ENERGY (J)	NOTCH RADIUS (mm)	SPECIMEN DIMENSIONS (mm)	POTENTIAL ENERGY (J)	COMMENTS
40	2024-T351	1.807	0.25	4 by 5	22	Complete Brittle Fracture
41	6061-T6	3.084	0.50	4 by 5	358	Compare to 26, 27, 28
42	6061-T6	3.084	0.50	4 by 5	358	Compare to 26, 27, 28
43	6061-T6	3.084	0.50	4 by 5	358	Compare to 26, 27, 28
44	6061-T6	3.168	0.25	4 by 5	358	Compare to A, B
C	7075	1.498	0.25	4 by 5	358	
D	7075	1.581	0.25	4 by 5	358	
E	7075	1.831	0.25	4 by 5	358	
F	7075	1.664	0.25	4 by 5	358	
45	4340	2.917	0.25	5 by 5	358	
46	4340	3.168	0.25		358	
47	4340	3.001	0.25	5 x 5	358	
48	4340	3.001	0.25	5 x 5	358	
49	4340	1.581	0.25	4 by 4	358	
56 - 72	4340-LL56	4.552	0.25	4 by 5	22	
56 - 73	4340-LL56	4.357	0.25	4 by 5	22	
56 - 74	4340-LL56	4.414	0.25	4 by 5	22	
56 - 1	4340-LL56	4.386	0.25	4 by 5	22	No side groove
56 - 2	4340-LL56	4.484	0.25	4 by 5	22	No side groove
56 - 3	4340-LL56	4.573	0.25	4 by 5	22	No side groove
56 - 4	4340-LL56	4.430	0.25	4 by 5	22	No side groove
56 - 5	4340-LL56	4.568	0.25	4 by 5	22	No side groove
56 -6	4340-LL56	4.470	0.25	4 by 5	22	No side groove
56 -7	4340-LL56	4.382	0.25	4 by 5	22	No side groove
56 -8	4340-LL56	4.457	0.25	4 by 5	22	No side groove
56 - 9	4340-LL56	4.546	0.25	4 by 5	22	No side groove
56 - 10	4340-LL56	4.298	0.25	4 by 5	22	No side groove
56 - 31S	4340-LL56	2.363	0.25	4 by 5	22	Side groove
56 - 32S	4340-LL56	2.467	0.25	4 by 5	22	Side groove
56 - 33S	4340-LL56	2.495	0.25	4 by 5	22	Side groove
56 - 34S	4340-LL56	2.519	0.25	4 by 5	22	Side groove
56 - 35S	4340-LL56	2.535	0.25	4 by 5	22	Side groove
56 - 36S	4340-LL56	2.483	0.25	4 by 5	22	Side groove
56 - 37S	4340-LL56	2.447	0.25	4 by 5	22	Side groove
56 - 38S	4340-LL56	2.535	0.25	4 by 5	22	Side groove
56 - 39S	4340-LL56	2.447	0.25	4 by 5	22	Side groove
56 - 40S	4340-LL56	2.404	0.25	4 by 5	22	Side groove
LL11-1	4340	2.888	0.25	4 by 5	22	Machine #2, no side groove

SPECIMEN ID	MATERIAL	ABSORBED ENERGY (J)	NOTCH RADIUS (mm)	SPECIMEN DIMENSIONS (mm)	POTENTIAL ENERGY (J)	COMMENTS
LL11-2	4340	2.739	0.25	4 by 5	22	Machine #2, no side groove
LL11-3	4340	2.725	0.25	4 by 5	22	Machine #2, no side groove
LL11-4	4340	2.793	0.25	4 by 5	22	Machine #2, no side groove
LL11-5	4340	2.725	0.25	4 by 5	22	Machine #2, no side groove
LL11-6	4340	2.684	0.25	4 by 5	22	Machine #2, no side groove
LL11-7	4340	2.698	0.25	4 by 5	22	Machine #2, no side groove
LL11-8	4340	2.671	0.25	4 by 5	22	Machine #2, no side groove
LL11-9	4340	2.671	0.25	4 by 5	22	Machine #2, no side groove
LL11-10	4340	2.752	0.25	4 by 5	22	Machine #2, no side groove
LA46-1	4340	2.210	0.25	4 by 5	22	Machine #2, side groove
LA46-2	4340	2.115	0.25	4 by 5	22	Machine #2, side groove
LA46-3	4340	1.424	0.25	4 by 5	22	Machine #2, side groove
LA46-4	4340	1.356	0.25	4 by 5	22	Machine #2, side groove
LA46-5	4340	1.315	0.25	4 by 5	22	Machine #2, side groove
LA46-6	4340	1.302	0.25	4 by 5	22	Machine #2, side groove
LA46-7	4340	1.342	0.25	4 by 5	22	Machine #2, side groove
LA46-8	4340	1.410	0.25	4 by 5	22	Machine #2, side groove
LA46-9	4340	1.356	0.25	4 by 5	22	Machine #2, side groove
LA46-10	4340	1.396	0.25	4 by 5	22	Machine #2, side groove
G7	2219-T87	0.759	0.25	4 by 5	22	Machine #2, no side groove
G8	2219-T87	0.705	0.25	4 by 5	22	Machine #2, no side groove
G9	2024T-351	1.071	0.25	4 by 5	22	Machine #2, no side groove
G10	2024T-351	1.071	0.25	4 by 5	22	Machine #2, no side groove

SPECIMEN ID	MATERIAL	ABSORBED ENERGY (J)	NOTCH RADIUS (mm)	SPECIMEN DIMENSIONS (mm)	POTENTIAL ENERGY (J)	COMMENTS
G13	2014-T6	2.386	0.25	4 by 5	22	Machine #2, no side groove
G14	2014-T6	1.302	0.25	4 by 5	22	Machine #2, no side groove
G3	7075-T6	1.152	0.25	4 by 5	22	Machine #2, no side groove
G4	7075-T6	0.936	0.25	4 by 5	22	Machine #2, no side groove
G1	6061-T6	2.088	0.50	4 by 5	22	Machine #2, no side groove
G2	6061-T6	1.966	0.50	4 by 5	22	Machine #2, no side groove
G5	6061-T6	1.790	0.50	4 by 5	22	Machine #2, no side groove
G6	6061-T6	2.766	0.50	4 by 5	22	Machine #2, no side groove
G11	7075-T6	0.664	0.25	4 by 5	22	Machine #2, no side groove
G12	7075-T6	0.610	0.25	4 by 5	22	Machine #2, no side groove

Table A-2. Data From Izod Tests.

SPECIMEN ID	MATERIAL	ABSORBED ENERGY (J)	NOTCH RADIUS (mm)	SPECIMEN DIMENSIONS (mm)	POTENTIAL ENERGY (J)	COMMENTS
1	4340	1.311	0.25	4 by 4	22	Remaining Ligament
2	4340	1.777	0.25	4 by 4	22	Remaining Ligament
3	6061-T6	2.743	0.50	4 by 5	22	Remaining Ligament
4	6061-T6	2.891	0.50	4 by 5	22	Remaining Ligament
5	6061-T6	2.580	0.50	4 by 5	22	Remaining Ligament
6	6061-T6	2.817	0.50	4 by 5	22	Remaining Ligament
7	6061-T6	2.461	0.50	4 by 5	22	Remaining Ligament
A	7075-T6	1.756	0.25	4 by 5	22	Remaining Ligament
B	7075-T6	1.390	0.25	4 by 5	22	Remaining Ligament
8	4340	1.331	0.25	4 by 4	22	Remaining Ligament
9	4340	1.314	0.25	4 by 4	22	Remaining Ligament
10	4340	1.314	0.25	4 by 4	22	Remaining Ligament
11	4340	2.043	0.25	4 by 5	22	Remaining Ligament
12	4340	2.150	0.25	4 by 5	22	Remaining Ligament
14	4340	2.125	0.25	4 by 5	22	Remaining Ligament
15	4340	2.484	0.25	5 by 5	22	Remaining Ligament
16	4340	2.449	0.25	5 by 5	22	Remaining Ligament
17	4340	2.557	0.25	5 by 5	22	Remaining Ligament
18	4340	2.537	0.25	5 by 5	22	Remaining Ligament
19	4340	2.731	0.25	5 by 5	22	Remaining Ligament

Table A-3. Combined Izod and Charpy Data.

IZOD

SPECIMEN ID.	MATERIAL	MEAN ABSORBED ENERGY (J)	STANDARD DEVIATION (J)	COEFFICIENT OF VARIATION	NOTCH RADIUS (mm)	SPECIMEN DIMENSION (mm)	POTENTIAL ENERGY (J)	COMMENTS
1-2, 8-10	4340	1.314	0.111	0.084	0.25	4 by 4		Remaining Ligament
34765	6061-T6	2.698	0.176	0.065	0.50	4 by 5		Remaining Ligament
A & B	7075-T6	1.573	0.259	0.165	0.50	4 by 5		Remaining Ligament
35017	4340	2.106	0.558	0.264	0.50	4 by 5		Remaining Ligament
15 - 19	4340	2.511	0.108	0.042	0.50	5 by 5		Remaining Ligament

CHARPY

SPECIMEN ID.	MATERIAL	MEAN ABSORBED ENERGY (J)	STANDARD DEVIATION (J)	COEFFICIENT OF VARIATION	NOTCH RADIUS (mm)	SPECIMEN DIMENSION (mm)	POTENTIAL ENERGY (J)	COMMENTS
20 - 22	6061-T6	3.521	0.162	0.046	0.50	4 by 5	5.0	Remaining Ligament (20 & 21) Complete Break - 22
23 - 25	6061-T6	3.474	0.150	0.043	0.50	4 by 5	22.0	Remaining Ligament (23 & 24) Complete Break - 25
26 - 28	6061-T6	3.110	0.300	0.097	0.25	4 by 5	22.0	Complete Breaks
29 - 31	7075-T6	1.172	0.058	0.050	0.25	4 by 5	22.0	Complete Laminate Breaks
32 - 34	2219-T87	1.316	0.069	0.052	0.25	4 by 5	22.0	Complete Brittle Fracture
SPECIMEN ID.	MATERIAL	MEAN ABSORBED ENERGY (J)	STANDARD DEVIATION (J)	COEFFICIENT OF VARIATION	NOTCH RADIUS (mm)	SPECIMEN DIMENSION (mm)	POTENTIAL ENERGY (J)	COMMENTS
35 - 37	2014-T6	2.894	0.105	0.036	0.25	4 by 5	22.0	Complete Laminated Fracture
38 - 40	2024-T351	0.838	0.088	0.047	0.25	4 by 5	22.0	Complete Brittle Fracture
41 - 43	6061-T6	3.084	0.000	0.000	0.50	4 by 5	358	Compare to 26, 27, 28, (44)
44	6061-T6	3.168	0.000	0.000	0.25	4 by 5	358	

C - F	7075	1.644	0.142	0.086	0.25	4 by 5	358	
45 - 48	4340	3.022	0.105	0.035	0.25	5 x 5	358	
56-72 - 56-74	4340-LL56	4.441	0.100	0.023	0.25	4 by 5	22	
56-1 - 56-10	4340-LL56	4.459	0.089	0.020	0.25	4 by 5	22	No side groove
56-31S - 56-40S	4340-LL56	2.476	0.068	0.028	0.25	4 by 5	22	Side groove
LL11-1 - LL11-10	4340	2.739	0.068	0.025	0.25	4 by 5	22	Machine #2, no side groove
LA46-1 - LA46-10	4340	1.519	0.339	0.223	0.25	4 by 5	22	Machine #2, side groove
G7 - G8	2219-T87	0.732	0.041	0.056	0.25	4 by 5	22	Machine #2, no side groove
G9 - G10	202T4-351	1.071	0.000	0.000	0.25	4 by 5	22	Machine #2, no side groove
G13 - G14	2014-T6	1.844	0.732	0.397	0.25	4 by 5	22	Machine #2, no side groove
G3 - G4	7075-T6	1.044	0.149	0.143	0.50	4 by 5	22	Machine #2, no side groove
G1 - G2	6061-T6	2.027	0.081	0.040	0.50	4 by 5	22	Machine #2, no side groove
G5 - G6	6061-T6	2.278	0.691	0.304	0.25	4 by 5	22	Machine #2, no side groove
G11 - G12	7075-T6	0.637	0.041	0.064	0.25	4 by 5	22	Machine #2, no side groove

Table 1
Effect of Alloy [Notes 1 to 3]

MATERIAL	NUMBER OF SPECIMENS	MEAN ABSORBED ENERGY (J)	STANDARD DEVIATION (J)	COEFFICIENT OF VARIATION
4340	3	4.441	0.100	0.023
2014-T6	3	2.894	0.105	0.036
2024-T351	3	1.868	0.088	0.047
2219-T87	3	1.316	0.069	0.052
6061-T6	3	3.110	0.300	0.097
7075-T6	3	1.172	0.058	0.050

Note 1: Notch Radius 0.25 mm
Note 2: Specimen Dimension 4 mm by 5mm
Note 3: Potential Energy 22 J

Table 2
Effect of Side Grooves [Notes 1 to 3]

MATERIAL	NUMBER OF SPECIMENS	MEAN ABSORBED ENERGY (J)	STANDARD DEVIATION (J)	COEFFICIENT OF VARIATION	SIDE GROOVES
4340	10	4.459	0.089	0.020	N
4340	10	2.476	0.068	0.028	Y

Note 1: Notch Radius 0.25 mm
Note 2: Specimen Dimension 4 mm by 5 mm
Note 3: Potential Energy 22 J

Table 3
Effect of Notch Radius [Notes 1 and 2]

MATERIAL	NUMBER OF SPECIMENS	MEAN ABSORBED ENERGY (J)	STANDARD DEVIATION (J)	COEFFICIENT OF VARIATION	NOTCH RADIUS (mm)
6061-T6	3	3.474	0.150	0.043	0.50
6061-T6	3	3.110	0.300	0.970	0.25

Note 1: Specimen Dimension 4 mm by 5 mm
Note 2: Potential Energy 22 J

Table 4
Effect of Machine Capacity [Notes 1 and 2]

MATERIAL	NUMBER OF SPECIMENS	MEAN ABSORBED ENERGY (J)	STANDARD DEVIATION (J)	COEFFICIENT OF VARIATION	NOTCH RADIUS (mm)	MACHINE POTENTIAL ENERGY (J)
6061-T6	3	3.521	0.162	0.046	0.50	5
6061-T6	3	3.474	0.150	0.043	0.50	22
6061-T6	3	3.110	0.300	0.097	0.25	22
6061-T6	3	3.084	0.000	0.000	0.50	358
6061-T6	1	3.168	0.000	0.000	0.25	358
7075-T6	4	1.644	0.142	0.086	0.25	358
7075-T6	3	1.172	0.058	0.050	0.25	22

Note 1: Specimen Dimension 4mm by 5 mm
Note 2: Potential Energy 22 J

C.N. McCowan,[1] M.J. Leap,[2] B. Gassner,[3] D.P. Vigliotti[1]

Standard and Experimental Heat Treatments for Low-Energy 4340 Steel Charpy V-notch Verification Specimens

Introduction

There is little information available on the production of the Charpy V-notch verification specimens used to certify impact machines to ASTM E 23 requirements. This is the first of several reports that will provide archival information on verification specimens for ASTM subcommittee E 28.07 (Impact Testing), which presides over the verification testing procedures. This report documents the results of several heat treatments for 4340 steels, used to produce low-energy impact specimens.

Summary

Five groups of specimens were heat-treated to evaluate the effect of different tempering conditions on the variability of absorbed energy. Three of the groups evaluated were determined to be of a quality suitable for use as verification specimens: groups 1a, 1b, and 3. The specimens from groups 1a and 1b were heat-treated according to recommendations given in ASTM E 1271, and the results show that impact specimens with very low standard deviations in absorbed energy can be produced using these recommendations. The group 3 specimens were heat-treated to promote the formation of fine carbides in the microstructure, and although they did qualify for use as verification specimens, the standard deviation in the absorbed energy for group 3 was higher than that for the group 1 specimens.

Two groups, 2a and 2b, varied too much in impact energy to be suitable for use as verification specimens. These specimens were tested only at room temperature, not −40 °C (−40 °F), and low-temperature testing would be expected to increase the scatter in the test results. Temper conditions used for groups 2a and 2b were designed to produce verification specimens for room-temperature testing. So, it appears that this approach is not too promising for use in the verification program.

---

[1] NIST, Materials Reliability Division, Boulder CO

[2] The Timken Company, Canton OH

[3] CNS CO. INC., Fullerton CA

Procedures

Machining

A total of 130 Charpy V-notch specimens were machined for the test matrix. Blanks were machined to near-final specimen dimensions for heat treatment. The specimens were machined to final size for nondestructive testing, then notched on the face labeled W in **Figure 1** and impact-tested.

Heat Treating

All the specimens were normalized at 900 °C (1650 °F) for 1 hour, air cooled, austenitized at 870 °C (1600 °F) for 1 hour, and oil-quenched prior to tempering treatments. The tempering treatments that were used are as follows:

(1) The specimens from group 1 were tempered according to the schedule given in ASTM E 1271: tempered at 399 °C (750 °F) for 1.5 hours, and quenched in oil. No specifications on the quench oil were cited. Two different companies produced group 1 specimens, and these specimens were labeled as group 1a and group 1b.

(2) The specimens from group 2 were tempered at 200 °C (392 °F) for 1 hour. Half of these specimens, group 2a, were tested in this condition, and the other half, group 2b, were cryogenically treated prior to testing. The cryogenic treatment consisted of slow cooling to −184 °C (−300 °F), followed by reheating to 182 °C (360 °F).

(3) The specimens for group 3 were given a two-stage tempering treatment. These specimens were tempered at 200 °C (392 °F) for 1 hour, then tempered at 399 °C (750 °F) for 1.5 hours.

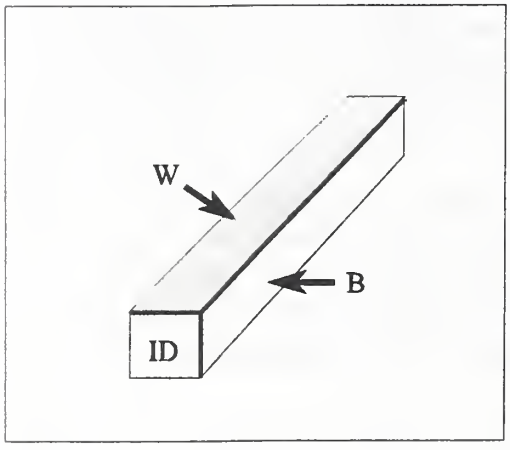


Figure 1. Specimen orientation.

Elastic Property Measurements

Elastic properties of the specimens were measured by impulse excitation of vibration as described in ASTM C1259. This is a nondestructive technique that consists of exciting the appropriate mode of vibration in the specimen by a single tap and determining the fundamental resonant frequency.

The resonant frequency of the unnotched impact specimens was determined in four modes of vibration: flexure and transverse flexure, and longitudinal and torsional vibration. These frequencies, along with the weight and dimensions of the specimens, were used to calculate elastic properties. Values of Young's modulus were calculated for flexural and longitudinal (E_f and E_l) vibration, as well as the torsion modulus, G, and Poisson's ratio. The raw frequency data, used to calculate

the elastic properties, are identified as RF, RL, and RT. These data vary with specimen dimension and weight.

For density calculation the dimensions of the specimens were individually measured to determine their width, breadth, and length. After the elastic properties were measured, notches were ground and the specimens were impact-tested.

Impact Testing

The five groups of specimens were impact-tested on a U-type impact machine equipped with an encoder and a digital display. The testing was done in accordance with ASTM E 23 procedures.

Groups 1a, 1b, and 3 were tested at −40 °C (−40 °F). Groups 2a and 2b were tested at room temperature.

Data Analysis

Data from the impact tests and the elastic-property measurements were evaluated with a commercially available statistical software program.

For the impact test results, the principal statistics of interest were the mean absorbed energy and the standard deviation (SD) in absorbed energy. The coefficient of variation (CV), or relative SD, which is the SD divided by the mean energy, is used as a normalized value to compare the variation in absorbed energies among the five groups of specimens.

For the moduli measurements, data were analyzed mainly to determine whether there was a correlation between the variation in absorbed energies and the variation in elastic measurements of the five specimen groups. The raw frequency data (RF, RL, and RT) were evaluated along with the elastic data to evaluate trends that might be indicative of variations in sample size.

Results and Discussion

Impact Tests

The CV values of the 5 groups of specimens are shown in **Figure 2**. Data for the specimens are given in the Appendix.

Groups 1a and 1b had the lowest CV values of the 5 groups tested. These values correspond to a mean energy of 18.4 J (13.6 ft-lbf) and a SD of 0.77 J (0.57 ft-lbf) for the 1a group, and a mean energy of 17.2 J (12.7 ft-lbf) and SD of 0.39 J (0.29 ft-lbf) for the 1b group.

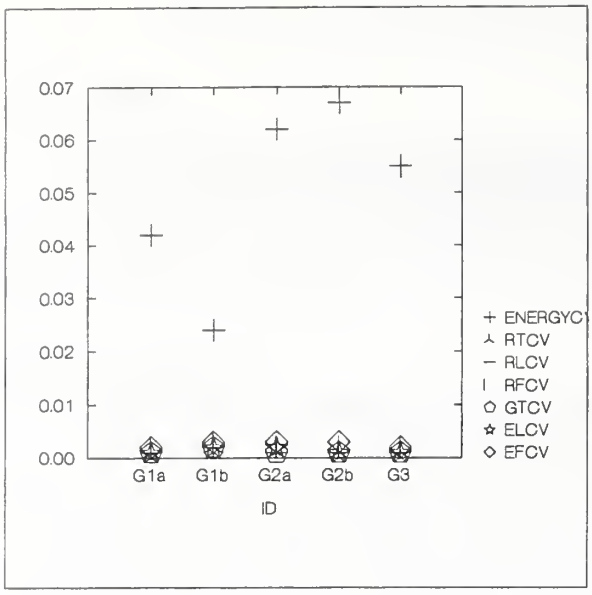


Figure 2. The coefficients of variation for the absorbed energy and the elastic property variables calculated for the five specimen groups

As described in Part I of this report, a SD of 1.02 J (0.75 ft-lbf) is the maximum allowed for low-energy verification specimens (12 to 20 J). So, both the 1a and 1b groups have low variations in absorbed energy. Group 1b has CV of 0.024, which is very low. The CV of group 1a, 0.04, is typical of CV values for verification specimens NIST has produced in the past.

Groups 2a and 2b have high CV values, 0.062 and 0.067 respectively. The SD for these 2 groups both exceeds the maximum of 1.0 J allowed by the NIST verification criteria. These specimens were tested at room temperature, which might increase the scatter in absorbed energy values compared with the other specimen groups tested here, but we did expect 2a would differ from 2b in the scatter due to the cryogenic temper treatment used for the 2b group. Little difference is apparent between these groups, and the specimens in group 2b, which were cryogenically tempered, had more scatter in absorbed energy than the specimens in group 2a.

The group 3 specimens have lower variation in absorbed energy than the 2a and 2b groups, but not as low as that for the 1a and 1b groups. The CV for group 3 was 0.055. This corresponds to a SD of 0.98 J (0.716 ft-lbf) and a mean energy of 17.6 J (13.0 ft-lbf). This group does meet the criteria for verification specimens. The low-temperature double temper was used for these specimens to determine whether the first temper would result in *seeding* fine carbides in the structure, resulting in a reduced scatter in absorbed energy. Apparently this approach shows little promise.

Elastic Properties

The values of CV for the six elastic constants measured for the sample groups are shown in **Figure 3**. If the CV values for the group 1b specimens were lower than those of the other groups, the general trend here would be similar to the trend in scatter for the absorbed energy (Figure 2,+ symbols). Group 1b, however, shows more scatter in elastic properties than do groups 1a and 3, and it had the lowest scatter in absorbed energy. So the elastic property measurements evaluated here provided little guidance in predicting which specimen groups would have more uniform impact properties.

Reviewing the raw frequency data in the Appendix, we observed that the CV values for the RF, RL, and RT are highest for the group 1b data. This may indicate that the final dimensions of these specimens (after notching) were less uniform than the other groups. These specimens were machined by a different shop than the one making the other specimens. Because the scatter in the frequency data is propagated through the elastic property calculations and finally to Poisson's ratio, the group 1b specimens may have artificially high scatter in elastic properties: one outlier was removed from the group 1b data, but several other points on the RF plots in the Appendix may also be outliers. The data for group 2a and 2b are interesting because these groups have more scatter in their elastic property data than do the other groups, but have a scatter in frequency data similar to that of the other groups: this may indicate measurement errors in specimen dimensions.

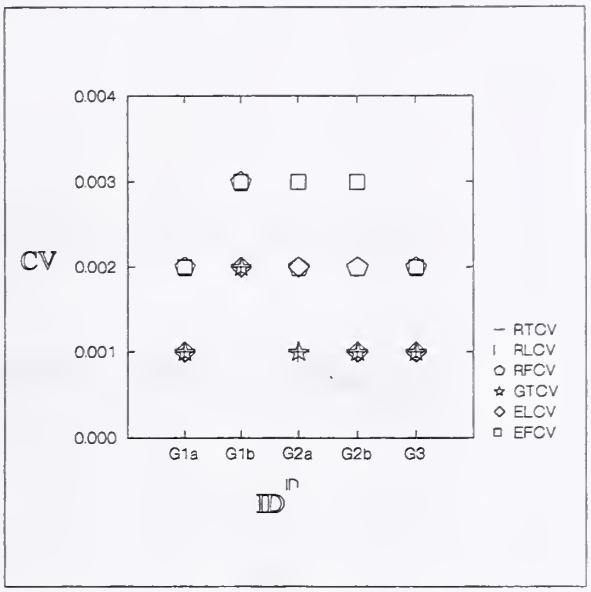


Figure 3. The coefficients of variation for the elastic properties are shown here on a more appropriate scale than in Figure 2.

The final observation apparent in the scatter plots in the Appendix is that there is not a strong correlation between impact energy of the specimens and the elastic property measurements. The range in impact energy of the group 1a, 1b, and 3 specimens, for example, probably represents a range in hardness of less than 1 HRC (from our experience with 4340 steel). Considering the summary data plotted in **Figure 4**, EL decreases as the energy increases (strength decreases).

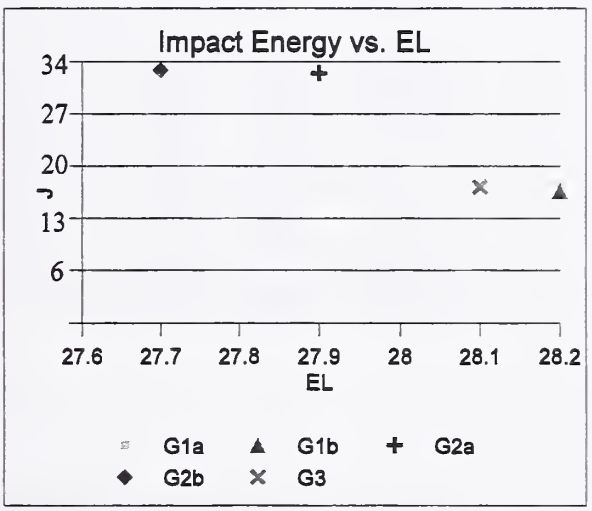


Figure 4: Impact energy (J) versus EL for the specimen groups.

We had hoped that the elastic property measurements would indicate which specimen groups have more homogeneous microstructures, and previous measurements have been shown to correlate well with hardness measurements. Given the localized notch used for the impact test, the dynamic nature of the test, and the very small differences in absorbed energies we are considering here, however, it is not too surprising that a better correlation between the scatter in absorbed energy and elastic properties was not found.

Appendix A: Elastic Property and Absorbed Energy Data

Summary Data

Group 1a (flexure)

	EF	EL	GT	RF	RL	RT	Energy
N OF CASES	25	25	25	25	25	25	25
MINIMUM	28.060	28.100	10.880	15.660	46.350	26.36	12.369
MAXIMUM	28.290	28.230	10.920	15.760	46.530	26.63	14.668
RANGE	0.230	0.130	0.040	0.100	0.180	0.090	2.299
MEAN	28.188	28.162	10.905	15.722	46.460	25.512	13.548
VARIANCE	0.003	0.001	0.000	0.001	0.002	0.001	0.330
STANDARD DEV	0.058	0.036	0.012	0.031	0.049	0.023	0.574
STD.ERROR	0.012	0.007	0.002	0.006	0.010	0.005	0.115
SKEWNESS (G1)	-0.436	0.045	-0.401	-0.600	-0.613	-0.332	0.099
KURTOSIS (G2)	-0.222	-1.027	-0.480	-0.854	-0.562	-0.010	-0.214
C.V.	0.002	0.001	0.001	0.002	0.001	0.001	0.042
MEDIAN	28.190	28.160	10.900	15.730	46.470	25.61	13.517

Group 1a (torsion)

	EF	EL	GT	RF	RL	RT
N OF CASES	25	25	25	25	25	25
MINIMUM	28.010	28.100	10.880	15.650	46.350	25.38
MAXIMUM	28.300	28.230	10.920	15.760	46.530	25.53
RANGE	0.290	0.130	0.040	0.110	0.180	0.09
MEAN	28.184	28.162	10.905	15.716	46.460	25.51
VARIANCE	0.005	0.001	0.000	0.001	0.002	0.001

Summary Data (continued)

Group 1a (torsion)

	EF	EL	GT	RF	RL	RT
STANDARDDEV	0.067	0.036	0.012	0.031	0.049	0.025
STD.ERROR	0.013	0.007	0.002	0.006	0.010	0.005
SKEWNESS (G1)	-0.793	0.045	-0.401	-0.511	-0.613	-0.332
KURTOSIS (G2)	0.478	-1.027	-0.480	-0.547	-0.562	-0.31
C.V.	0.002	0.001	0.001	0.002	0.001	0.001
MEDIAN	28.190	28.160	10.900	15.720	46.470	25.51

Group 1b (flexure)

	EF	EL	GT	RF	RL	RT	Energy
N OF CASES	30	30	30	30	30	30	30
MINIMUM	27.820	28.050	10.860	15.640	46.330	26.540	11.987
MAXIMUM	28.370	28.370	10.990	25.670	46.660	26.730	13.325
RANGE	0.550	0.320	0.130	10.030	0.330	0.190	1.338
MEAN	28.236	28.201	10.923	16.038	46.425	26.599	12.692
VARIANCE	0.014	0.003	0.000	3.312	0.006	0.002	0.090
STANDARD DEV	0.082	0.056	0.022	1.820	0.074	0.043	0.293
STD.ERROR	0.022	0.010	0.004	0.332	0.014	0.008	0.055
SKEWNESS (G1)	-1.396	0.159	0.099	5.194	1.162	0.959	0.054
KURTOSIS (G2)	3.032	2.184	2.782	24.999	1.520	1.016	-0.154
C.V.	0.003	0.002	0.002	0.003	0.002	0.002	0.024
MEDIAN	28.255	28.210	10.920	15.705	46.410	26.590	12.687

Summary Data (continued)

Group1b (torsion)

	EF	EL	GT	RF	RL	RT
N OF CASES	30	30	30	30	30	30
MINIMUM	28.070	28.050	10.860	15.630	46.330	26.540
MAXIMUM	28.420	28.370	10.990	15.830	46.660	26.730
RANGE	0.350	0.320	0.130	0.200	0.330	0.190
MEAN	28.225	28.201	10.923	15.687	46.425	26.599
VARIANCE	0.005	0.003	0.000	0.002	0.006	0.002
STANDARD DEV	0.073	0.056	0.022	0.045	0.074	0.043
STD.ERROR	0.013	0.010	0.004	0.008	0.014	0.008
SKEWNESS (G1)	0.035	0.177	0.102	1.283	1.162	0.959
KURTOSIS (G2)	0.535	2.220	3.094	1.740	1.520	1.016
C.V.	0.003	0.002	0.002	0.003	0.002	0.002
MEDIAN	28.225	28.210	10.920	15.675	46.410	26.590

Group 2a (flexure)

	EF	EL	GT	RF	RL	RT	Energy
N OF CASES	25	25	25	25	25	25	25
MINIMUM	27.810	27.790	10.740	15.550	46.100	26.370	21.915
MAXIMUM	28.090	27.990	10.810	15.640	46.240	26.470	26.571
RANGE	0.280	0.200	0.070	0.090	0.140	0.100	4.656
MEAN	27.927	27.874	10.774	15.584	46.155	26.412	23.957
VARIANCE	0.007	0.002	0.000	0.001	0.002	0.001	2.209
STANDARD DEV	0.084	0.044	0.016	0.024	0.039	0.025	1.486

Summary Data (continued)

Group 2a (flexure)

	EF	EL	GT	RF	RL	RT	Energy
STD.ERROR	0.017	0.009	0.003	0.005	0.008	0.005	0.297
SKEWNESS (G1)	0.144	0.489	0.016	0.681	0.647	0.532	0.713
KURTOSIS (G2)	-1.031	0.329	0.092	-0.460	-0.531	-0.148	-0.979
C.V.	0.003	0.002	0.001	0.002	0.001	0.001	0.062
MEDIAN	27.950	27.870	10.770	15.580	46.150	26.410	23.287

Group 2a (torsion)

	EF	EL	GT	RF	RL	RT
N OF CASES	25	25	25	25	25	25
MINIMUM	27.810	27.790	10.740	15.550	46.100	26.370
MAXIMUM	28.170	27.990	10.810	15.640	46.240	26.470
RANGE	0.360	0.200	0.070	0.090	0.140	0.100
MEAN	27.932	27.874	10.774	15.589	46.155	26.412
VARIANCE	0.010	0.002	0.000	0.001	0.002	0.001
STANDAR DDEV	0.098	0.044	0.016	0.026	0.039	0.025
STD.ERROR	0.020	0.009	0.003	0.005	0.008	0.005
SKEWNESS (G1)	0.544	0.489	0.016	0.657	0.647	0.532
KURTOSIS (G2)	-0.495	0.329	0.092	-0.380	-0.531	-0.148
SUM	698.310	696.860	269.360	389.720	1153.87	660.310
C.V.	0.004	0.002	0.001	0.002	0.001	0.001
MEDIAN	27.930	27.870	10.770	15.580	46.150	26.410

Summary Data (continued)

Group 2b (flexure)

	EF	EL	GT	RF	RL	RT	Energy
N OF CASES	25	25	25	25	25	25	25
MINIMUM	27.470	27.640	10.680	15.540	46.090	26.370	21.534
MAXIMUM	27.750	27.780	10.730	15.630	46.240	26.460	26.967
RANGE	0.280	0.140	0.050	0.090	0.150	0.090	5.433
MEAN	27.612	27.708	10.703	15.590	46.161	26.410	24.236
VARIANCE	0.007	0.001	0.000	0.001	0.002	0.001	2.601
STANDARDDEV	0.085	0.037	0.014	0.028	0.046	0.028	1.613
STD.ERROR	0.017	0.007	0.003	0.006	0.009	0.006	0.323
SKEWNESS (G1)	-0.010	-0.085	0.417	-0.019	0.254	0.278	0.231
KURTOSIS (G2)	-1.047	-0.320	-0.599	-1.288	-1.225	-1.301	-1.226
C.V.	0.003	0.001	0.001	0.002	0.001	0.001	0.067
MEDIAN	27.610	27.710	10.700	15.590	46.150	26.400	23.680

Group 2b (torsion)

	EF	EL	GT	RF	RL	RT
N OF CASES	25	25	25	25	25	25
MINIMUM	27.400	27.640	10.680	15.540	46.090	26.370
MAXIMUM	27.770	27.780	10.730	15.640	46.240	26.460
RANGE	0.370	0.140	0.050	0.100	0.150	0.090
MEAN	27.584	27.709	10.703	15.586	46.162	26.410
VARIANCE	0.007	0.001	0.000	0.001	0.002	0.001
STANDARDDEV	0.081	0.038	0.014	0.029	0.046	0.028

Summary Data (continued)

Group 2b (torsion)

	EF	EL	GT	RF	RL	RT
STD.ERROR	0.016	0.008	0.003	0.006	0.009	0.006
SKEWNESS (G1)	-0.282	-0.123	0.385	0.472	0.220	0.321
KURTOSIS (G2)	0.940	-0.423	-0.510	-0.777	-1.191	-1.273
C.V.	0.003	0.001	0.001	0.002	0.001	0.001
MEDIAN	27.580	27.710	10.700	15.580	46.150	26.400

Group 3 (flexure)

	EF	EL	GT	RF	RL	RT	Energy
N OF CASES	25	25	25	25	25	25	25
MINIMUM	27.890	28.020	10.860	15.670	46.390	26.560	11.162
MAXIMUM	28.150	28.210	10.910	15.760	46.510	26.650	14.284
RANGE	0.260	0.190	0.050	0.090	0.120	0.090	3.122
MEAN	28.045	28.096	10.877	15.716	46.442	26.602	13.004
VARIANCE	0.004	0.002	0.000	0.001	0.001	0.000	0.512
STANDARD DEV	0.064	0.041	0.013	0.025	0.038	0.022	0.716
STD.ERROR	0.013	0.008	0.003	0.005	0.008	0.004	0.143
SKEWNESS (G1)	-0.347	0.560	0.724	-0.176	0.175	-0.017	-0.307
KURTOSIS (G2)	-0.148	0.818	0.256	-1.026	-1.164	-0.505	0.396
C.V.	0.002	0.001	0.001	0.002	0.001	0.001	0.055
MEDIAN	28.050	28.090	10.880	15.720	46.440	26.600	13.006

Summary Data (continued)

Group3 (torsion)

	EF	EL	GT	RF	RL	RT
N OF CASES	25	25	25	25	25	25
MINIMUM	27.970	28.020	10.860	15.680	46.390	26.560
MAXIMUM	28.160	28.210	10.910	15.770	46.510	26.650
RANGE	0.190	0.190	0.050	0.090	0.120	0.090
MEAN	28.049	28.096	10.877	15.720	46.442	26.602
VARIANCE	0.003	0.002	0.000	0.001	0.001	0.000
STANDARDDEV	0.056	0.041	0.013	0.024	0.038	0.022
STD.ERROR	0.011	0.008	0.003	0.005	0.008	0.004
SKEWNESS (G1)	0.293	0.560	0.724	0.087	0.175	-0.017
KURTOSIS (G2)	-0.920	0.818	0.256	-0.786	-1.164	-0.505
C.V.	0.002	0.001	0.001	0.002	0.001	0.001
MEDIAN	28.050	28.090	10.880	15.720	46.440	26.600

Raw Frequency Data and Elastic Property Data

Group 1a

ID$	RF	RL	RT	EF	EL	GT	ENERGY
N7	15.72	46.47	26.61	28.19	28.18	10.90	14.22
N43	15.75	46.51	26.63	28.29	28.23	10.92	14.54
N10	15.70	46.44	26.59	28.26	28.22	10.92	14.66
N41	15.67	46.39	26.56	28.22	28.18	10.90	12.94
N12	15.74	46.48	26.63	28.18	28.14	10.90	12.36
N71	15.71	46.45	26.61	28.21	28.17	10.91	13.58
N19	15.76	46.53	26.64	28.24	28.20	10.91	14.41
N78	15.75	46.51	26.64	28.23	28.19	10.91	12.81
N20	15.68	46.39	26.59	28.18	28.15	10.92	13.64
N86	15.73	46.47	26.63	28.16	28.13	10.90	13.70
N22	15.72	46.42	26.60	28.06	28.10	10.89	13.38
N44	15.75	46.50	26.64	28.18	28.13	10.90	13.13
N29	15.72	46.47	26.61	28.16	28.16	10.90	12.56
N39	15.72	46.45	26.60	28.20	28.19	10.91	14.28
N30	15.74	46.47	26.63	28.10	28.12	10.90	13.26
N46	15.76	46.51	26.65	28.17	28.12	10.90	13.45
N34	15.73	46.46	26.61	28.09	28.13	10.89	13.32
N52	15.70	46.43	26.59	28.15	28.12	10.88	13.58
N37	15.66	46.37	26.56	28.19	28.16	10.90	13.51
N60	15.76	46.51	26.63	28.26	28.19	10.91	13.90
N61	15.67	46.35	26.56	28.21	28.15	10.91	13.45
N62	15.75	46.48	26.64	28.27	28.18	10.92	13.77
N69	15.68	46.41	26.58	28.20	28.20	10.92	13.32
N45	15.75	46.51	26.64	28.10	28.11	10.88	13.32

Raw Frequency Data and Elastic Property Data (continued)

Group 1a

ID$	RF	RL	RT	EF	EL	GT	ENERGY
N67	15.74	46.51	26.64	28.21	28.19	10.92	13.51
T34	15.72	46.46	26.61	28.18	28.13	10.89	13.32
T37	15.66	46.37	26.56	28.15	28.16	10.90	13.51
T29	15.72	46.47	26.61	28.20	28.16	10.90	12.56
T30	15.73	46.47	26.63	28.16	28.12	10.90	13.26
T45	15.75	46.51	26.64	28.01	28.11	10.88	13.32
T39	15.71	46.45	26.60	28.25	28.19	10.91	14.28
T46	15.76	46.51	26.65	28.17	28.12	10.90	13.45
T10	15.70	46.44	26.59	28.26	28.22	10.92	14.66
T52	15.70	46.43	26.59	28.06	28.12	10.88	13.58
T12	15.74	46.48	26.63	28.18	28.14	10.90	12.39
T60	15.72	46.51	26.63	28.26	28.19	10.91	13.90
T19	15.76	46.53	26.64	28.24	28.20	10.91	14.41
T61	15.65	46.35	26.56	28.17	28.15	10.91	13.45
T20	15.69	46.39	26.59	28.22	28.15	10.92	13.64
T62	15.73	46.48	26.64	28.19	28.18	10.92	13.77
T22	15.69	46.42	26.60	28.07	28.10	10.89	13.38
T67	15.74	46.51	26.64	28.21	28.19	10.92	13.51
T44	15.75	46.50	26.64	28.14	28.13	10.90	13.13
T69	15.68	46.41	26.58	28.30	28.20	10.92	13.32
T41	15.67	46.39	26.56	28.18	28.18	10.90	12.94
T71	15.71	46.45	26.61	28.21	28.17	10.91	13.58
T7	15.72	46.47	26.61	28.19	28.18	10.90	14.22
T78	15.75	46.51	26.64	28.23	28.19	10.91	12.81

Raw Frequency Data and Elastic Property Data (continued)

Group 1a

ID$	RF	RL	RT	EF	EL	GT	ENERGY
T43	15.74	46.51	26.63	28.25	28.23	10.92	14.54
T86	15.72	46.47	26.63	28.12	28.13	10.90	13.71

Group1b

ID$	RF	RL	RT	EF	EL	GT	ENERGY
NA1	15.71	46.38	26.60	28.21	28.16	10.93	12.49
NF5	15.64	46.34	26.54	28.16	28.21	10.92	12.30
NA2	15.67	46.35	26.56	28.16	28.15	10.91	12.87
NF7	15.69	46.39	26.59	28.14	28.15	10.91	13.32
NA3	15.68	46.39	26.57	28.25	28.21	10.92	12.36
NF6	15.67	46.36	26.56	28.16	28.14	10.90	12.94
NA4	15.73	46.47	26.61	28.30	28.23	10.92	12.49
NF4	15.67	46.35	26.56	28.14	28.16	10.91	12.94
NA5	15.78	46.52	26.66	28.24	28.18	10.92	12.30
NJ6	15.67	46.36	26.57	28.33	28.23	10.94	13.00
NA6	15.65	46.33	26.54	28.31	28.21	10.93	12.56
NJ7	15.65	46.34	26.55	28.28	28.23	10.94	12.75
NA7	15.69	46.41	26.58	28.37	28.24	10.93	12.56
NJ8	15.72	46.44	26.61	28.36	28.24	10.94	11.98
NA8	15.72	46.46	26.62	28.13	28.18	10.92	13.00
NJ9	15.75	46.49	26.64	28.19	28.19	10.92	12.68
NA9	15.67	46.46	26.62	28.34	28.37	10.99	12.68
NJ10	15.69	46.38	26.58	28.36	28.24	10.94	13.19
NA10	15.72	46.42	26.61	28.31	28.20	10.94	12.75
NF8	15.71	46.43	26.61	28.12	28.12	10.90	12.62

Raw Frequency Data and Elastic Property Data (continued)

Group1b

IDS	RF	RL	RT	EF	EL	GT	ENERGY
NF1	15.73	46.47	26.63	28.09	28.14	10.90	12.68
NJ5	15.71	46.43	26.59	28.27	28.22	10.92	12.36
NF2	15.65	46.37	26.55	28.20	28.18	10.90	12.81
NF9	15.67	46.39	26.56	28.17	28.21	10.91	12.36
NF10	15.76	46.50	26.65	28.36	28.22	10.94	12.75
NJ1	15.86	46.66	26.73	28.37	28.28	10.95	12.56
NJ4	15.69	46.41	26.59	28.26	28.22	10.93	12.56
NJ2	15.75	46.50	26.63	28.34	28.23	10.93	12.62
NJ3	15.70	46.40	26.58	28.33	28.24	10.93	13.19
TF2	15.67	46.37	26.55	28.15	28.18	10.90	12.81
TF3	15.77	46.56	26.67	28.07	28.05	10.86	12.94
TF1	15.71	46.47	26.63	28.14	28.14	10.90	12.68
TA4	15.71	46.47	26.61	28.26	28.23	10.92	12.49
TA5	15.75	46.52	26.66	28.21	28.18	10.92	12.30
TF9	15.66	46.39	26.56	28.26	28.21	10.91	12.36
TA6	15.63	46.33	26.54	28.27	28.21	10.93	12.56
TF10	15.74	46.50	26.65	28.28	28.22	10.94	12.75
TA7	15.67	46.41	26.58	28.28	28.24	10.93	12.56
TJ1	15.83	46.66	26.73	28.42	28.28	10.94	12.56
TA8	15.70	46.46	26.62	28.19	28.18	10.92	13.00
TJ2	15.73	46.50	26.63	28.22	28.23	10.93	12.62
TA9	15.66	46.46	26.62	28.30	28.37	10.99	12.68
TJ3	15.68	46.40	26.58	28.25	28.24	10.93	13.19
TA10	15.69	46.42	26.61	28.23	28.20	10.94	12.75

Raw Frequency Data and Elastic Property Data (continued)

Group1b

ID$	RF	RL	RT	EF	EL	GT	ENERGY
TJ4	15.67	46.41	26.59	28.22	28.22	10.93	12.56
TF8	15.69	46.43	26.61	28.08	28.12	10.90	12.62
TJ5	15.68	46.43	26.59	28.15	28.21	10.92	12.36
TF4	15.65	46.35	26.50	28.19	28.16	10.91	12.94
TJ6	15.65	46.36	26.57	28.28	28.23	10.94	13.00
TA1	15.68	46.38	26.60	28.22	28.16	10.93	12.49

Raw Frequency Data and Elastic Property Data (continued)

Group 2a

ID$	RF	RL	RT	EF	EL	GT	ENERGY
N1	15.56	46.13	26.40	27.84	27.85	10.77	22.96
N58	15.57	46.13	26.41	27.83	27.82	10.76	23.15
N5	15.57	46.13	26.40	28.04	27.94	10.80	23.09
N57	15.57	46.12	26.40	28.01	27.89	10.78	23.74
N11	15.59	46.16	26.41	27.96	27.87	10.77	23.02
N91	15.60	46.17	26.42	27.99	27.87	10.77	22.76
N14	15.56	46.11	26.40	27.96	27.88	10.79	24.59
N96	15.57	46.15	26.42	27.83	27.85	10.77	26.30
N21	15.56	46.11	26.38	27.81	27.82	10.75	23.28
N97	15.61	46.19	26.44	27.81	27.82	10.76	22.30
N24	15.58	46.14	26.40	28.00	27.87	10.77	25.64
N65	15.56	46.12	26.38	28.07	27.93	10.79	22.69
N27	15.55	46.10	26.38	27.84	27.86	10.77	23.28
N55	15.62	46.22	26.44	27.90	27.88	10.77	26.43
N35	15.56	46.11	26.37	27.86	27.85	10.75	23.41
N73	15.57	46.15	26.40	28.09	27.99	10.81	25.45
N47	15.58	46.15	26.41	27.96	27.91	10.79	22.76
N77	15.64	46.24	26.46	27.97	27.89	10.78	22.69
N51	15.60	46.18	26.44	27.89	27.85	10.77	23.48
N83	15.63	46.23	26.47	27.88	27.83	10.77	26.57
N84	15.60	46.18	26.42	27.97	27.92	10.79	26.57
N88	15.61	46.20	26.44	27.98	27.88	10.78	23.15
N90	15.59	46.14	26.41	27.81	27.79	10.74	23.68
N70	15.59	46.18	26.41	27.92	27.92	10.78	25.91

Raw Frequency Data and Elastic Property Data (continued)

Group 2a

ID$	RF	RL	RT	EF	EL	GT	ENERGY
N87	15.57	46.13	26.40	27.95	27.88	10.78	21.91
T47	15.58	46.15	26.41	28.05	27.91	10.79	22.76
T51	15.61	46.18	26.44	27.89	27.85	10.77	23.48
T27	15.55	46.10	26.38	27.97	27.86	10.77	23.28
T35	15.55	46.11	26.37	27.81	27.85	10.75	23.41
T70	15.60	46.18	26.41	28.05	27.92	10.78	25.91
T55	15.63	46.22	26.44	27.94	27.88	10.77	26.43
T73	15.58	46.15	26.40	28.17	27.99	10.81	25.45
T5	15.58	46.13	26.40	28.08	27.94	10.80	23.09
T77	15.64	46.24	26.46	27.97	27.89	10.78	22.69
T11	15.58	46.16	26.41	27.83	27.87	10.77	23.02
T83	15.64	46.23	26.47	27.83	27.83	10.77	26.57
T14	15.57	46.11	26.40	28.00	27.88	10.79	24.69
T84	15.60	46.18	26.42	28.02	27.92	10.79	26.57
T21	15.57	46.11	26.38	27.85	27.82	10.75	23.28
T88	15.63	46.20	26.44	27.93	27.88	10.78	23.15
T24	15.58	46.14	26.40	27.87	27.87	10.77	25.64
T87	15.57	46.13	26.40	27.95	27.88	10.78	21.91
T65	15.56	46.12	26.38	28.03	27.93	10.79	22.69
T90	15.59	46.14	26.41	27.81	27.79	10.74	23.68
T57	15.57	46.12	26.40	27.97	27.89	10.78	23.74
T91	15.59	46.17	26.42	27.82	27.87	10.77	22.76
T1	15.58	46.13	26.40	27.92	27.85	10.77	22.96
T96	15.58	46.15	26.42	27.87	27.85	10.77	26.30

Raw Frequency Data and Elastic Property Data (continued)

Group 2a

ID$	RF	RL	RT	EF	EL	GT	ENERGY
T58	15.58	46.13	26.41	27.83	27.82	10.76	23.15
T97	15.61	46.19	26.44	27.85	27.82	10.76	22.30

Raw Frequency Data and Elastic Property Data (continued)

Group 2b

ID$	RF	RL	RT	EF	EL	GT	ENERGY
N2	15.63	46.22	26.44	27.73	27.78	10.73	26.11
N56	15.58	46.15	26.40	27.61	27.72	10.70	26.24
N3	15.57	46.11	26.38	27.65	27.71	10.70	23.35
N50	15.56	46.10	26.38	27.59	27.68	10.70	22.17
N4	15.59	46.17	26.40	27.63	27.73	10.70	25.05
N93	15.62	46.20	26.44	27.69	27.72	10.71	26.50
N13	15.61	46.19	26.42	27.62	27.71	10.70	25.45
N95	15.59	46.14	26.40	27.49	27.64	10.68	25.25
N15	15.59	46.15	26.40	27.65	27.72	10.71	23.22
N99	15.62	46.18	26.43	27.69	27.72	10.72	23.54
N16	15.59	46.15	26.41	27.63	27.70	10.70	23.28
N59	15.62	46.22	26.46	27.47	27.64	10.69	23.02
N17	15.57	46.13	26.38	27.55	27.70	10.69	26.17
N42	15.62	46.23	26.45	27.60	27.71	10.70	23.61
N25	15.57	46.13	26.39	27.56	27.69	10.70	23.28
N64	15.63	46.24	26.45	27.73	27.76	10.72	23.68
N28	15.56	46.13	26.38	27.56	27.71	10.69	22.63
N66	15.56	46.12	26.39	27.57	27.70	10.70	22.30
N40	15.56	46.13	26.39	27.60	27.72	10.70	26.96
N74	15.58	46.12	26.40	27.49	27.65	10.69	24.07
N81	15.63	46.21	26.45	27.74	27.73	10.72	23.74
N82	15.60	46.18	26.43	27.75	27.77	10.73	26.70
N92	15.55	46.10	26.37	27.50	27.68	10.69	21.52
N63	15.62	46.23	26.44	27.70	27.76	10.72	25.51

Raw Frequency Data and Elastic Property Data (continued)

N89	15.54	46.09	26.37	27.50	27.66	10.68	22.43
T28	15.55	46.13	26.39	27.58	27.71	10.70	22.63
T40	15.58	46.15	26.39	27.64	27.74	10.70	26.96
T17	15.56	46.13	26.38	27.51	27.70	10.69	26.17
T25	15.57	46.13	26.39	27.56	27.69	10.70	23.28
T53	15.63	46.23	26.44	27.64	27.76	10.72	25.51
T42	15.63	46.23	26.45	27.62	27.71	10.70	23.61
T64	15.64	46.24	26.45	27.77	27.76	10.72	23.68
T3	15.55	46.11	26.38	27.58	27.71	10.70	23.35
T66	15.57	46.12	26.39	27.60	27.70	10.70	22.30
T4	15.58	46.17	26.40	27.60	27.73	10.70	25.05
T74	15.56	46.12	26.40	27.40	27.65	10.69	24.07
T13	15.60	46.19	26.42	27.58	27.71	10.70	25.45
T81	15.61	46.21	26.45	27.65	27.73	10.72	23.74
T15	15.59	46.15	26.40	27.65	27.72	10.71	23.22
T82	15.59	46.18	26.43	27.71	27.77	10.73	26.70
T16	15.58	46.15	26.40	27.61	27.70	10.70	23.28
T89	15.56	46.09	26.37	27.55	27.66	10.68	22.43
T59	15.64	46.22	26.46	27.55	27.64	10.69	23.02
T92	15.56	46.10	26.37	27.53	27.68	10.69	21.53
T50	15.54	46.10	26.38	27.52	27.68	10.70	22.17
T93	15.60	46.20	26.44	27.56	27.72	10.71	26.50
T2	15.61	46.22	26.44	27.65	27.78	10.73	26.11
T95	15.57	46.14	26.40	27.40	27.64	10.68	25.25
T56	15.57	46.15	26.40	27.57	27.72	10.70	26.24
T99	15.60	46.18	26.43	27.56	27.72	10.72	23.54

Raw Frequency Data and Elastic Property Data (continued)

Group 3

ID$	RF	RL	RT	EF	EL	GT	ENERGY
N6	15.72	46.43	26.60	28.04	28.07	10.87	13.58
N49	15.68	46.41	26.58	27.96	28.07	10.86	11.16
N8	15.71	46.42	26.60	27.98	28.04	10.86	13.00
N48	15.68	46.40	26.56	28.05	28.10	10.87	13.07
N9	15.74	46.49	26.62	28.15	28.21	10.91	12.87
N94	15.70	46.40	26.60	27.96	28.04	10.88	12.81
N18	15.74	46.48	26.61	28.08	28.13	10.88	13.58
N98	15.70	46.39	26.59	27.97	28.02	10.86	13.26
N23	15.73	46.47	26.61	28.12	28.15	10.89	13.45
N100	15.75	46.49	26.62	28.12	28.12	10.88	13.13
N26	15.68	46.39	26.57	28.06	28.10	10.88	13.13
N53	15.73	46.47	26.62	28.02	28.08	10.87	13.26
N31	15.71	46.43	26.59	28.05	28.09	10.87	12.94
N38	15.73	46.45	26.61	28.15	28.15	10.90	12.36
N32	15.72	46.45	26.61	28.09	28.09	10.88	11.98
N68	15.73	46.46	26.62	28.05	28.10	10.88	13.51
N33	15.67	46.43	26.59	27.89	28.05	10.86	12.36
N72	15.75	46.51	26.63	28.05	28.13	10.88	14.22
N36	15.69	46.39	26.57	28.09	28.10	10.88	12.87
N75	15.70	46.41	26.57	28.04	28.09	10.86	12.36
N76	15.72	46.44	26.61	28.01	28.09	10.88	12.17
N79	15.73	46.47	26.60	28.05	28.11	10.87	12.62
N85	15.74	46.47	26.63	28.12	28.13	10.90	14.15
N54	15.69	46.40	26.58	28.04	28.09	10.88	12.87

Raw Frequency Data and Elastic Property Data (continued)

Group 3

IDS	RF	RL	RT	EF	EL	GT	ENERGY
N80	15.76	46.51	26.65	27.99	28.06	10.87	14.28
T33	15.71	46.43	26.59	27.99	28.05	10.86	12.36
T36	15.69	46.39	26.57	28.06	28.10	10.88	12.87
T31	15.71	46.43	26.59	28.05	28.09	10.87	12.94
T32	15.72	46.45	26.61	28.01	28.09	10.88	11.98
T54	15.69	46.40	26.58	28.04	28.09	10.88	12.87
T38	15.72	46.45	26.61	28.10	28.15	10.90	12.36
T68	15.74	46.46	26.62	28.07	28.10	10.88	13.51
T8	15.71	46.42	26.60	27.98	28.04	10.86	13.00
T72	15.77	46.51	26.63	28.13	28.13	10.88	14.22
T9	15.73	46.49	26.62	28.15	28.21	10.91	12.87
T75	15.69	46.41	26.57	28.00	28.09	10.86	12.36
T18	15.75	46.48	26.61	28.16	28.13	10.88	13.58
T76	15.74	46.44	26.61	28.10	28.09	10.88	12.17
T23	15.73	46.47	26.61	28.09	28.15	10.89	13.45
T79	15.74	46.47	26.60	28.09	28.11	10.87	12.62
T26	15.68	46.39	26.57	28.04	28.10	10.88	13.13
T80	15.76	46.51	26.65	27.97	28.06	10.87	14.28
T53	15.73	46.47	26.62	28.00	28.08	10.87	13.26
T85	15.74	46.47	26.63	28.09	28.13	10.90	14.15
T48	15.68	46.40	26.56	28.02	28.10	10.87	13.07
T94	15.70	46.40	26.60	27.99	28.04	10.88	12.81
T6	15.72	46.43	26.60	27.99	28.07	10.87	13.58
T98	15.70	46.39	26.59	27.97	28.02	10.86	13.26

Raw Frequency Data and Elastic Property Data (continued)

Group 3

ID$	RF	RL	RT	EF	EL	GT	ENERGY
T49	15.71	46.41	26.58	28.06	28.07	10.86	11.16
T100	15.74	46.49	26.62	28.07	28.12	10.88	13.13

Summary Statistics and Data for the Initial CVN Specimens Evaluated

LL49

	EF	EL	GT	RF	ENERGY
N OF CASES	5.00	5.00	5.00	5.00	5.00
MINIMUM	25.40	28.05	11.06	14.57	95.79
MAXIMUM	25.51	28.15	11.09	14.60	100.98
RANGE	0.11	0.10	0.03	0.03	5.20
MEAN	25.45	28.09	11.07	14.58	98.64
VARIANCE	0.00	0.00	0.00	0.00	4.36
STANDARD DEV	0.04	0.04	0.01	0.01	2.09
STD.ERROR	0.02	0.02	0.01	0.01	0.93
SKEWNESS (G1)	0.11	1.09	0.87	0.27	-0.32
KURTOSIS (G2)	-1.14	-0.17	-0.27	-1.04	-1.31
C.V.	0.00	0.00	0.00	0.00	0.02
MEDIAN	25.44	28.08	11.07	14.58	99.19

LL55

	EF	EL	GT	RF	ENERGY
N OF CASES	5.00	5.00	5.00	5.00	5.00
MINIMUM	24.91	27.74	10.91	14.43	93.33
MAXIMUM	25.12	27.84	10.95	14.49	102.12
RANGE	0.21	0.10	0.04	0.06	8.79
MEAN	24.99	27.78	10.92	14.45	96.92

Summary Statistics and Data for the Initial CVN Specimens Evaluated (continued)

LL55

	EF	EL	GT	RF	ENERGY
VARIANCE	0.01	0.00	0.00	0.00	13.09
STANDARD DEV	0.08	0.04	0.02	0.02	3.62
STD.ERROR	0.04	0.02	0.01	0.01	1.62
SKEWNESS (G1)	0.80	0.84	1.29	0.85	0.35
KURTOSIS (G2)	-0.82	-0.76	-0.08	-0.73	-1.14
C.V.	0.00	0.00	0.00	0.00	0.04
MEDIAN	24.95	27.76	10.91	14.44	97.58

HH54

	EF	EL	GT	RF	ENERGY
N OF CASES	5.00	5.00	5.00	5.00	5.00
MINIMUM	25.43	28.25	11.13	14.58	17.20
MAXIMUM	25.57	28.38	11.17	14.62	18.76
RANGE	0.14	0.13	0.04	0.04	1.56
MEAN	25.52	28.32	11.15	14.60	17.84
VARIANCE	0.00	0.00	0.00	0.00	0.38
STANDARD DEV	0.05	0.05	0.02	0.02	0.62
STD.ERROR	0.02	0.02	0.01	0.01	0.28
SKEWNESS (G1)	-0.80	-0.35	-0.37	-0.75	0.56
KURTOSIS (G2)	-0.67	-0.50	-0.78	-0.64	-1.08

Summary Statistics and Data for the Initial CVN Specimens Evaluated (continued)

HH54

	EF	EL	GT	RF	ENERGY
C.V.	0.00	0.00	0.00	0.00	0.04
MEDIAN	25.54	28.32	11.15	14.61	17.63

HH55

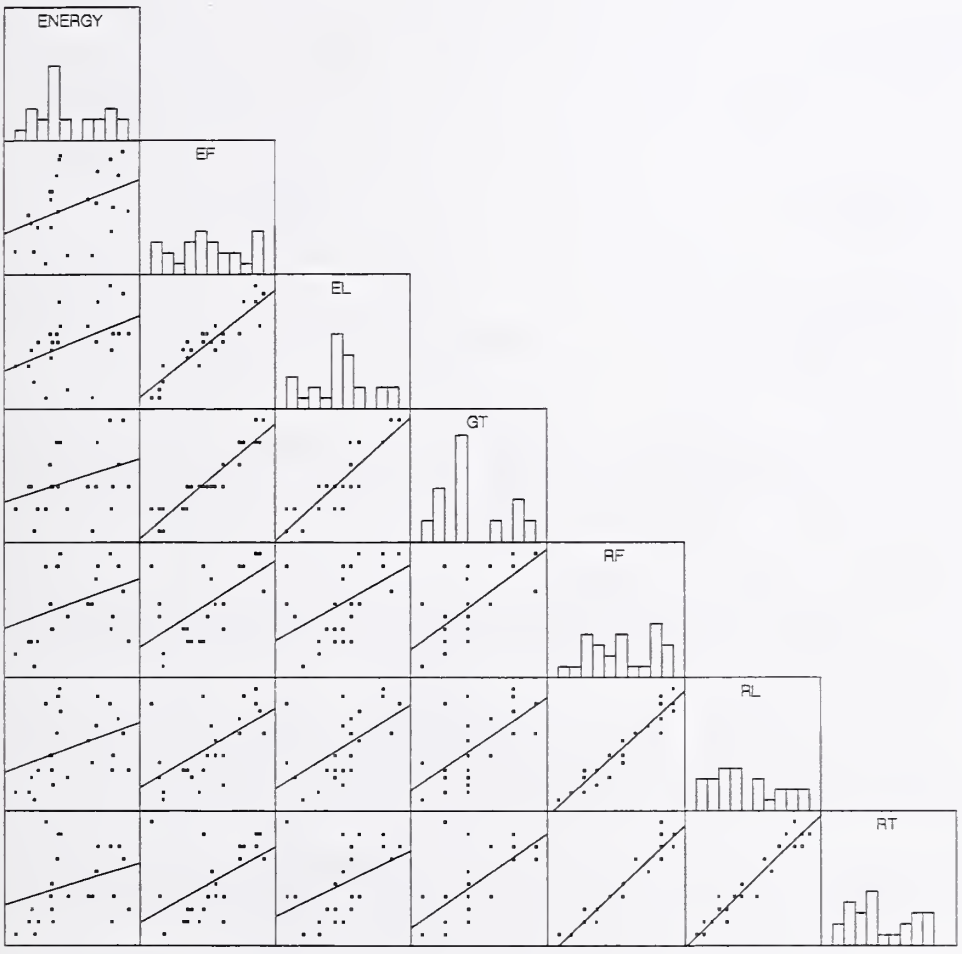
	EF	EL	GT	RF	ENERGY
N OF CASES	5.00	5.00	5.00	5.00	5.00
MINIMUM	25.36	28.18	11.10	14.56	18.24
MAXIMUM	25.54	28.27	11.14	14.61	19.97
RANGE	0.18	0.09	0.04	0.05	1.73
MEAN	25.44	28.22	11.12	14.58	19.21
VARIANCE	0.01	0.00	0.00	0.00	0.53
STANDARD DEV	0.08	0.04	0.02	0.02	0.73
STD.ERROR	0.04	0.02	0.01	0.01	0.33
SKEWNESS (G1)	0.33	0.09	0.35	0.38	-0.15
KURTOSIS (G2)	-1.62	-1.41	-1.42	-1.59	-1.39
C.V	0.00	0.00	0.00	0.00	0.04
MEDIAN	25.40	28.22	11.11	14.57	19.11

Data for First Pilot-Lot Series

CODE$	RF	RL	RT	EF	EL	GT	ENERGY
LL49	14.57	45.09	26.06	25.40	28.05	11.06	97.30
LL49	14.59	45.12	26.08	25.48	28.08	11.07	99.19
LL49	14.60	45.17	26.10	25.51	28.15	11.09	99.94
LL49	14.58	45.12	26.07	25.44	28.08	11.07	95.79
LL49	14.58	45.11	26.07	25.44	28.07	11.07	100.98
LL55	14.43	44.84	25.88	24.91	27.74	10.91	93.61
LL55	14.44	44.86	25.89	24.94	27.76	10.91	97.96
LL55	14.46	44.88	25.90	25.02	27.79	10.92	93.33
LL55	14.44	44.85	25.88	24.95	27.75	10.91	102.12
LL55	14.49	44.92	25.93	25.12	27.84	10.95	97.58
HH54	14.61	45.31	26.17	25.54	28.32	11.15	17.46
HH54	14.58	45.25	26.15	25.43	28.25	11.13	17.63
HH54	14.62	45.36	26.19	25.57	28.38	11.17	18.15
HH54	14.61	45.32	26.18	25.54	28.33	11.16	18.76
HH54	14.60	45.31	26.17	25.50	28.32	11.15	17.20
HH55	14.57	45.20	26.12	25.40	28.18	11.11	18.24
HH55	14.60	45.25	26.14	25.51	28.25	11.13	19.97
HH55	14.56	45.21	26.11	25.36	28.20	11.10	19.11
HH55	14.61	45.27	26.16	25.54	28.27	11.14	18.85
HH55	14.57	45.23	26.12	25.40	28.22	11.11	19.89

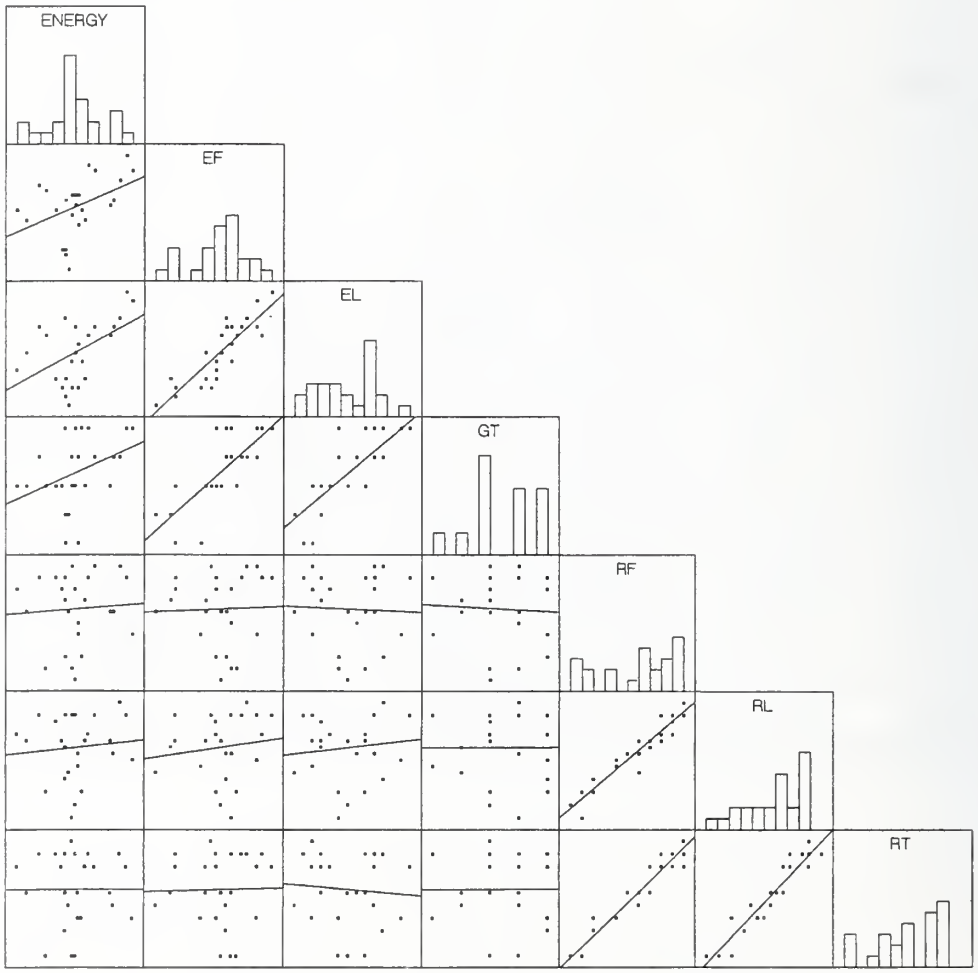
Scatter Plots for the Absorbed Energy, Elastic Property, and Frequency Data

Group 2B Scatter Plots



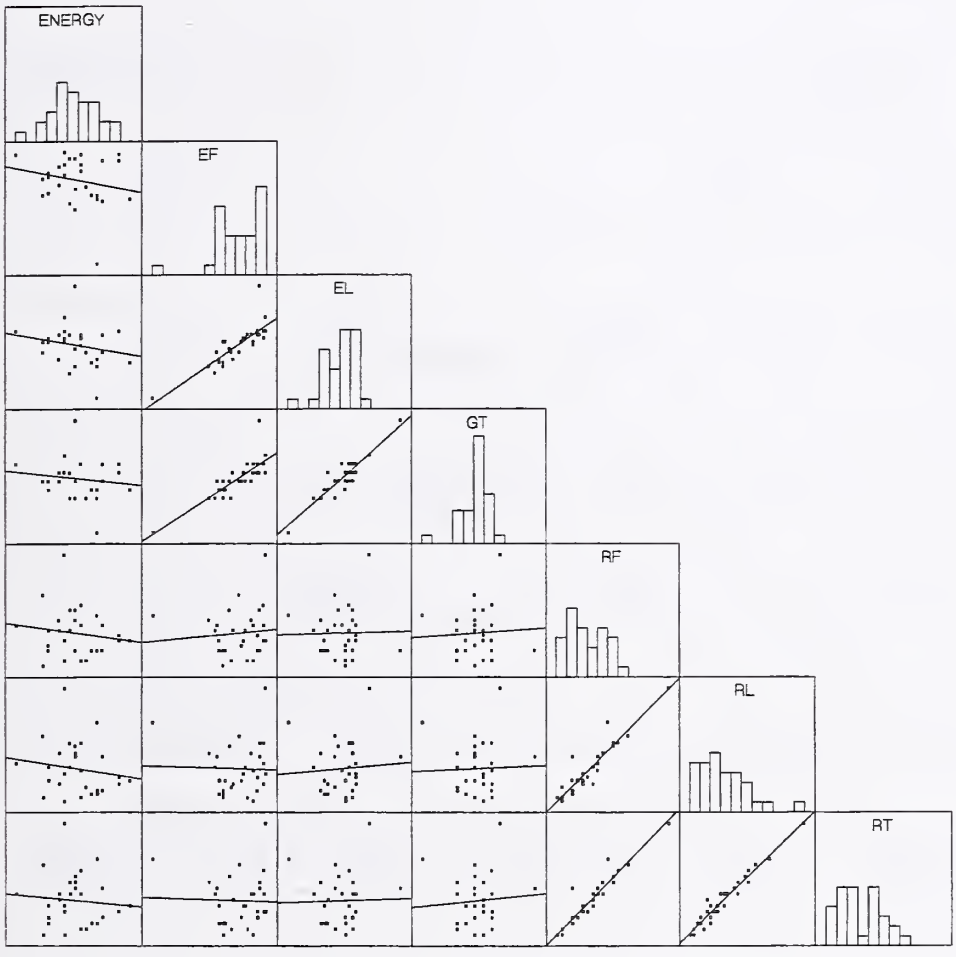
Scatter Plots for Specimen Groups

Group 1a Scatter Plot (flexure mode)



Scatter Plots for Specimen Groups (continued)

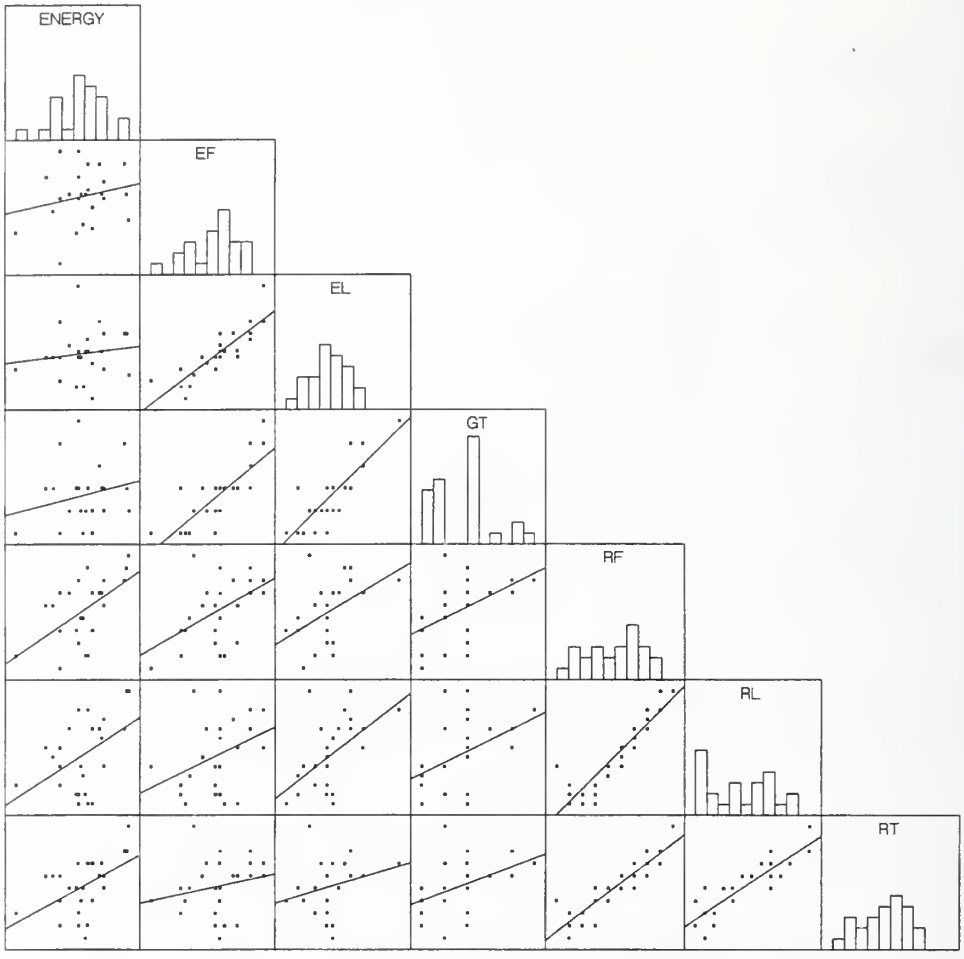
Group 1b Scatter Plot[4] (torsional mode)



[4] The outlier, apparent in the EF-EL scatter plot was removed from the data, and the summary statistics. The point was identified as nf3.

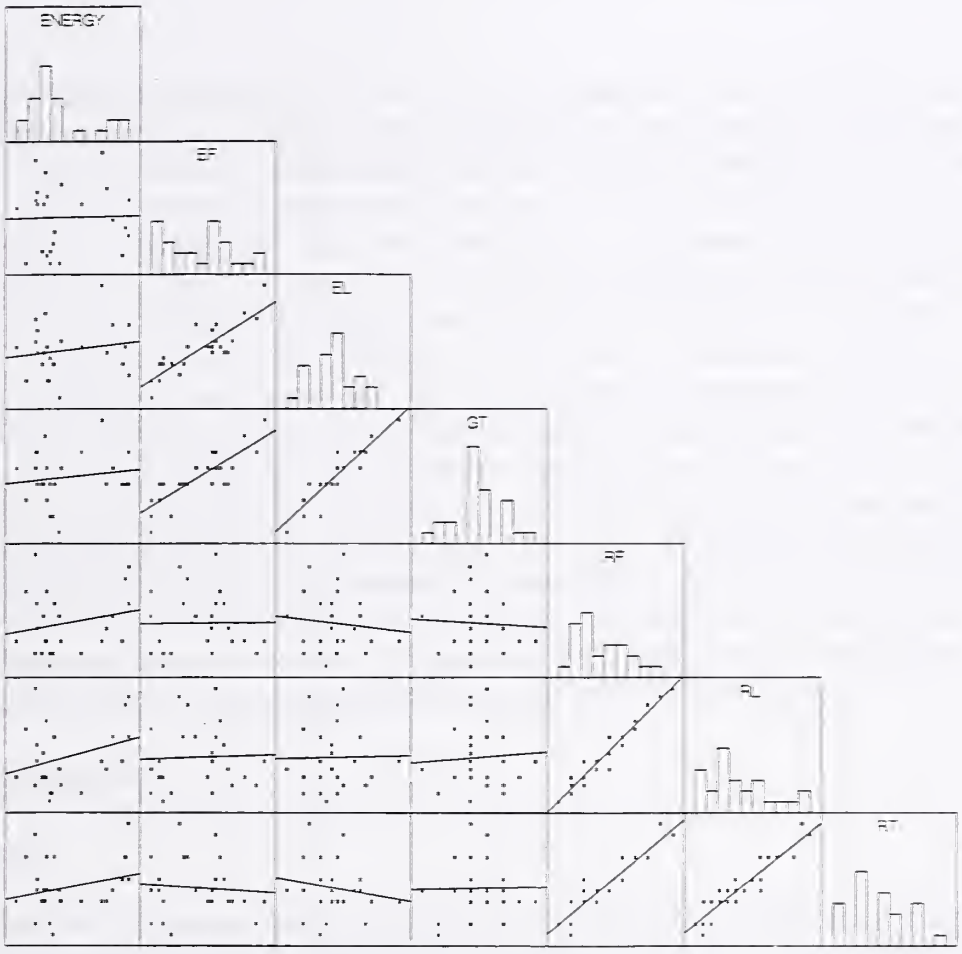
Scatter Plots for Specimen Groups (continued)

Group 3 Scatter Plot (flexural mode)



Scatter Plots for Specimen Groups (continued)

Group 2a Scatter Plot (flexural mode)



C.N. McCowan[1] and D.P. Vigliotti[1]

Homogenization and Subcritical Annealing of 4340 Charpy V-notch Verification Specimens

Introduction

When considering the possible contributors to scatter in the impact properties of 4340 quenched and tempered Charpy verification specimens, inhomogeneities in the as-received microstructure of the material are always a concern. In the Charpy Impact Verification Program at NIST, impact verification specimens are produced from hot-rolled bar stock. The bar is rolled from a single-melt ingot of VIM-VAR 4340 steel. We track both the ingot and the general location within the ingot from which each bar is made. By not mixing bar stock from different ingot locations when making a production lot of verification specimens and by tightly controlling the chemistry of the VIM-VAR steel, we minimize these contributions to inhomogeneity as much as possible. However, as the hot-rolled bars are stacked for cooling during the rolling operation, it would be expected that individual bars are exposed to somewhat different cooling conditions. If so, then carbide precipitation may vary from bar to bar, and these carbides affect the transgranular fracture resistance of the steel. Large carbides are not expected to fully dissolve during subsequent heat treatments, which typically did not exceed 900 °C. Recent work has shown that modification of these residual carbide networks during heat treatment between 800 and 900 °C can result in significant differences in the optimum toughness and the scatter in toughness for these steels. So, in a continuing effort to minimize the scatter in impact toughness of the Charpy verification specimens, the possible benefit of re-homogenizing our 4340 steel was investigated.

Materials and Procedures

Sample Preparation

The Charpy V-notch verification specimens were made from VIM-VAR 4340 steel, with low sulfur and phosphorus contents (0.00 % S & 0.004 % P; residual V content is approximately 0.04 mass %). We purchased the steel as hot-rolled square bar. The bar was cut to length and ground to near sample-size dimensions (10 mm × 10 mm × 60 mm) prior to heat treating. Following heat treatment, specimens were machined to final dimensions and the notches were ground. The specimens used in this study were randomly selected from a production lot that had been prepared for heat treating. They should contain a representative sample of the bar stock.

[1]NIST Materials Reliability Division, Boulder, CO

Heat Treatments

Two groups of specimens, Groups A and B, were heat-treated for this study. Each group contained 30 specimens. The Group A specimens were homogenized and subcritically annealed prior to hardening and tempering. The Group B specimens, the control group, were annealed prior to hardening and tempering following our standard practice. The two groups were combined together in one basket for the hardening and tempering treatments to minimize differences due to these steps in the heat treatment. Specifics of the heat treatment are given below.

The Group A specimens were homogenized at 1100 °C (2012 °F) for 75 min in a vacuum furnace and cooled with nitrogen gas at one atmosphere. The rate at which the gas-quench cooled the samples can be approximated by two linear segments: (1) cooling from 1100 to 816 °C (2012 to 1500 °F) took 40 min, which represents a cooling rate of near 7 °C/min (12 °F/min), and (2) cooling from 816 to 38 °C (1500 to 100 °F) took 50 min, which represents a cooling rate near 16 °C/min (27 °F/min). The homogenized samples were then subcritically annealed at 650 °C (1200 °F) for 60 min to seed the microstructure with fine carbides (cooled at a rate of 17.3 °C/min).

The Group B samples were normalized in a vacuum furnace at 893 °C (1640 °F) for 90 min and gas cooled at a rate of 18.6 °C/min.

Groups A and B were combined for the hardening and tempering, and processed as follows: (1) held at 870 °C (1600 °F) for 100 min, (2) quenched in oil, (3) tempered at 588 °C (1090 °F) for 120 min in vacuum furnace and, (4) cooled at rate of 16.7 °C/min.

Mechanical Testing

The hardness of each specimen was measured by taking the average of two hardness values. These measurements were made on the face of the specimen that is opposite the notch, approximately 5 mm from the ends of the samples.

The impact tests were conducted at −40 °C (−40 °F). The impact machine used for the tests is equipped with an optical encoder and digital display that outputs the energy absorbed during the test.

Metallographic Samples

Three samples were taken from the groups at various stages during the heat treatment process; (1) a sample from Group A after the homogenization treatment, (2) a sample from Group A after the subcritical anneal, and (3) a sample from Group B after the normalizing. These samples were prepared for evaluation by light microscopy. In addition, samples were chosen following impact testing for evaluations of fracture surface in a scanning electron microscope.

Results and Discussion

Mechanical Test Data

The data for absorbed energy of the two groups of specimens are shown in **Figure 1** and listed in **Table 1**. Clearly, the different heat treatments used for the two groups prior to hardening and tempering treatments affected the distribution and average values of the absorbed energy data. The distribution of the Group A data is extremely skewed and the mean absorbed energy (68.4 J, 50.5 ft-lbf) for the Group A specimens is higher than that of the Group B specimens (61.6 J, 45.4 ft-lbf).

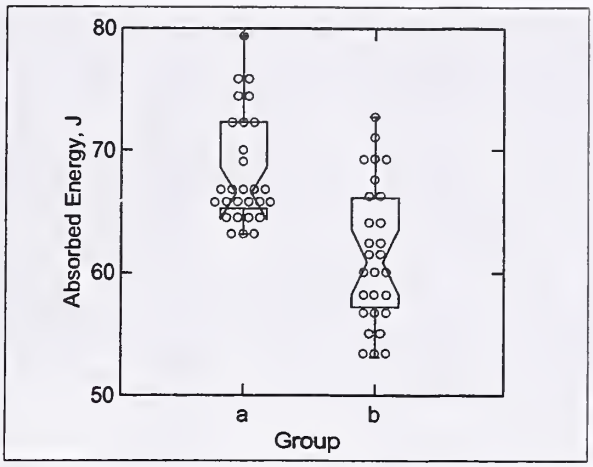


Figure 1. Absorbed energy data for Groups A and B, notched at the median values and boxed at upper and lower confidence levels of 95 %.

The range and standard deviation in the absorbed energy for the Group A data (16.2 J and 4.5 J) are lower than those for the Group B data (19.6 J and 5.7 J), indicating that the homogenization and subcritical annealing treatments helped to reduce the variation in absorbed energy in these tests. However, the skewed distribution of the absorbed energy data for Group A is troubling.

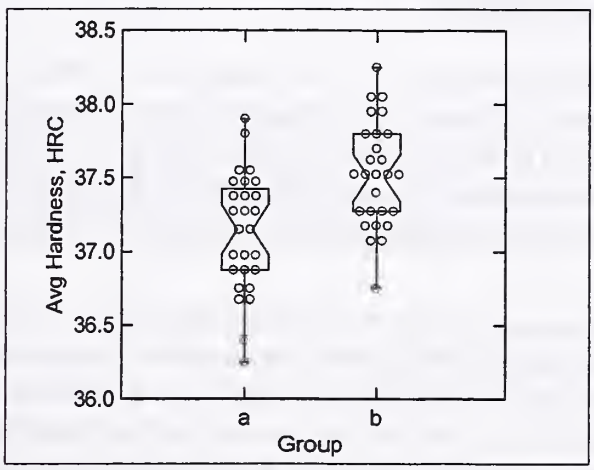


Figure 2. Hardness data for Groups A & B, notched at the median values and boxed at upper and lower confidence levels of 95 %.

The hardness data in **Figure 2** show fairly normal distributions for both groups of specimens. The variation for the Group A data (range of 1.7, standard deviation of 0.40 HRC) is slightly greater than that for the Group B data (range of 1.5, standard deviation of 0.36 HRC) , but not significantly. The mean hardness of Group A (37.2 HRC) is lower than that of Group B (37.5 HRC), as might be expected from the absorbed energy results, since hardness generally decreases as absorbed energy increases. However, within the small range in hardness for this study, the relationship between hardness and absorbed energy (**Figure 3**) is not too useful for predicting impact toughness: at a hardness of 37.5 HRC, the absorbed energy for the Group B specimens has a range of more than 15 J.

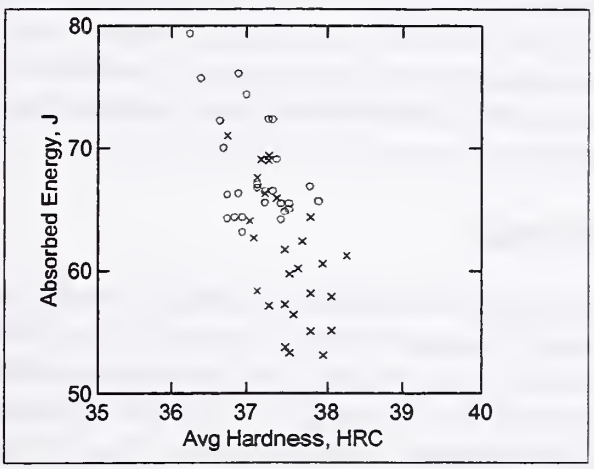


Figure 3. Absorbed energy versus hardness for the two specimen groups where "o" indicates Group A and "x" indicates Group B.

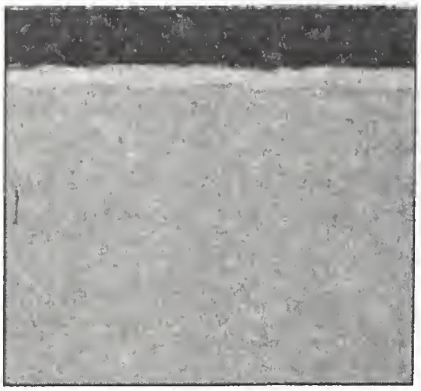


Figure 4. The decarburized layer (top) at the surface of the homogenized sample. Bar equal to 50 μm.

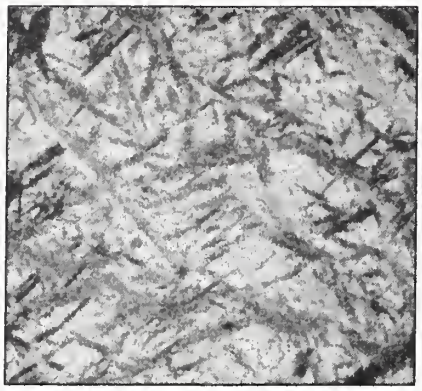


Figure 5. Lightly etched (picric acid) microstructure of the homogenized sample. Bar equal to 50 μm.

Figure 6. Microstructure of the homogenized and subcritically annealed sample. Bar equal to 50 μm.

Microstructure

The decarburization of the homogenized samples was approximately twice that of the control samples (Group B), but the depth of the decarburization should not present any problems. As shown in **Figure 4**, the decarburization is approximately 25 μm (0.002 in) deep for the homogenized sample. Very light machining can remove this surface. The interface between the decarburized layer and the matrix is smooth and free of stress raisers.

The grain size for the homogenized sample was quite large, as expected (200 to 300 μm). The microstructure, **Figure 5**, is presumed from optical-microscopy evaluations to be a combination of martensite and bainite. The balance of these constituents apparently changes across the thickness of the sample: near the surface, the darker acicular constituents dominate the microstructure. Microhardness measurements showed the dark region to have a hardness around 400 to 450 HV, and the matrix hardness to be around 500 to 550 HV. Following the subcritical annealing, the hardness through the sample was found to be more uniform. The average hardness of the subcritically annealed sample was around 375 HV. As shown in **Figure 6**, the microstructure of the subcritically annealed sample retains many remnants of the homogenized microstructure, but there appears to be a substantial amount of precipitation and speroidization of carbides throughout the structure, which results in the darker appearance of the matrix and a softening of the acicular features.

Fractography

Six samples were chosen to determine whether the differences in absorbed energy values for the Group A specimens could be related to fracture-surface characteristics. Three samples from Group A with low energy and three samples with high energy were used for this comparison. Even with the unaided eye, the fracture surfaces of the low- and high-energy samples were observed to differ: the color (shade of grey) of the higher-energy samples was characteristically darker than that of the lower-energy samples. However, at higher magnifications, no characteristic differences between the samples were observed. An example of the ductile

appearance of the fracture surfaces examined for these samples is shown in **Figure 7.**

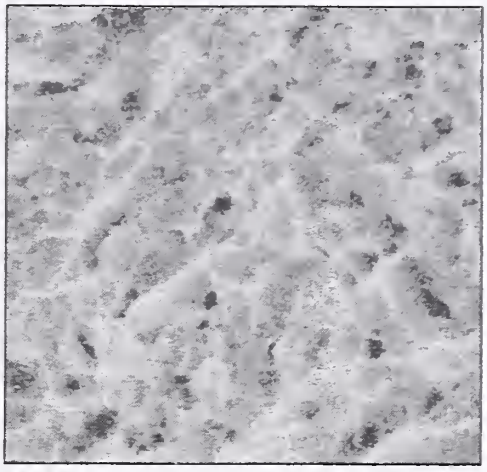


Figure 7. The typical appearance of the fracture surface near the notch of the Charpy specimens. The length of the bottom of the image is approximately 60 μm.

Summary

The comparison of the data for the two groups of specimens considered here is inconclusive. It is not clear that the homogenization and subcritical annealing of the Group A specimens significantly reduced the scatter in absorbed energy. However, the Group A data do have a lower standard deviation in absorbed energy than that of the Group B data, and the skewed distribution of the absorbed energy data for Group A contains 18 specimens that have very low scatter when considered separately (range of 4 J and standard deviation of 1.1 J). This offers some hope that a homogenization treatment can reduce the scatter in the verification specimens.

The homogenization treatment did not produce stress raisers on the surface of the samples, which would increase their susceptibility to quench cracking during hardening. So machining the specimens prior to hardening should not be necessary. The depth of decarburization was not excessive, so our machining practices (0.4 mm per side oversize) before heat treatment should be sufficient.

Since the results of this study show some promise of providing a more robust heat treatment for the impact verification specimens, the next step should be to conduct a comprehensive study that will include samples having a range of precipitated carbide. Most of the small carbides precipitated during a subcritical annealing treatment would be expected to dissolve during the hardening treatment, but a homogeneous distribution of carbides might be beneficial to reducing the scatter in impact energy. The point of the study should be to determine the optimum size and distribution of carbides required to reduce scatter in absorbed energy.

For any follow-up studies it is recommended that several subcritical annealing conditions be included in the test matrix to investigate the effect of size and distribution of carbides on impact properties, and that several homogenization treatments be investigated.

Table 1: Data for the Group A and B specimens

Case	GROUP	J	HR	HI	HAVG
1	a	75.698	36.200	36.600	36.400
2	a	79.347	36.300	36.200	36.250
3	a	72.247	36.900	36.400	36.650
4	a	65.557	37.100	37.400	37.250
5	a	70.012	36.700	36.700	36.700
6	a	67.040	37.200	37.100	37.150
7	a	74.391	37.000	37.000	37.000
8	a	65.463	37.500	37.600	37.550
9	a	66.483	37.600	36.900	37.250
10	a	64.167	37.300	37.600	37.450
11	a	65.001	37.600	37.500	37.550
12	a	63.150	37.000	36.900	36.950
13	a	64.353	36.700	37.200	36.950
14	a	64.353	36.800	36.900	36.850
15	a	72.340	37.600	37.000	37.300
16	a	64.815	37.700	37.300	37.500
17	a	66.761	36.900	37.400	37.150
18	a	76.073	36.900	36.900	36.900
19	a	72.340	37.500	37.200	37.350
20	a	65.649	38.000	37.800	37.900
21	a	65.463	37.500	37.400	37.450
22	a	66.205	36.900	36.600	36.750
23	a	69.082	37.400	37.400	37.400
24	a	66.483	37.400	37.300	37.350
25	a	66.297	36.800	37.000	36.900
26	a	66.854	37.900	37.700	37.800
27	a	64.259	36.900	36.600	36.750
28	a	75.513	.	.	.
29	b	67.597	37.100	37.200	37.150
30	b	58.171	37.800	37.800	37.800
31	b	60.566	38.100	37.800	37.950
32	b	59.736	37.800	37.300	37.550
33	b	57.895	38.200	37.900	38.050
34	b	60.197	37.700	37.600	37.650

Case	GROUP	J	HR	HI	HAVG
35	b	69.084	37.300	37.100	37.200
36	b	61.211	38.400	38.100	38.250
37	b	64.076	37.200	36.900	37.050
38	b	68.991	37.100	37.500	37.300
39	b	71.037	36.800	36.700	36.750
40	b	62.689	37.300	36.900	37.100
41	b	53.767	37.500	37.500	37.500
42	b	58.355	37.200	37.100	37.150
43	b	56.425	37.600	37.600	37.600
44	b	64.353	38.300	37.300	37.800
45	b	61.765	37.400	37.600	37.500
46	b	53.309	37.600	37.500	37.550
47	b	57.160	37.100	37.500	37.300
48	b	65.928	37.600	37.200	37.400
49	b	62.412	37.900	37.500	37.700
50	b	55.141	38.300	37.800	38.050
51	b	57.252	37.800	37.200	37.500
52	b	66.299	37.500	37.000	37.250
53	b	69.362	37.400	37.200	37.300
54	b	53.126	37.900	38.000	37.950
55	b	55.049	38.100	37.500	37.800
56	b	72.714	.	.	.

C. N. McCowan, T. A. Siewert, and D. P. Vigliotti

Heat Treatment of NIST T-200 CVN Specimens

Introduction

A T-type maraging steel was used to produce very-high-energy (near 200 J) verification specimens that are tested to certify the performance of Charpy impact test machines. The steel is an 18 % nickel alloy in which titanium (rather than cobalt) is used as the primary strengthening element. In a peak aged condition, these alloys would be expected to have Charpy impact energies of around 110 J (80 ft-lbs) and hardness of about 43 to 47 HRC (Rockwell C scale). NIST uses the alloy to produce specimens with much higher impact energy, however, and refers to these specimens as superhigh-energy specimens. Our super high energy verification specimens have impact energies typically in the range of 175 to 245 J (130 to 180 ft-lbs).

We recently purchased a new heat of T-type maraging steel and planned a study to help optimize the heat treatment for this new steel and reduce the variation in impact energy of the specimens. Issues of primary interest to this study include: (1) redistribution of indigenous inclusions by solution heat treatment, to evaluate the effect on the fracture energy and scatter in impact energy; (2) grain refinement and beneficial effects of multiple recrystallizations on the variation in absorbed energy: (3) grain size and morphology effects, and (4) controlled cooling and its effect on the degree of embrittlement and impact toughness.

Literature Review

There has been a significant amount of research done on T-type maraging steels, but there is still disagreement on several of the factors that interest us. It is beyond our scope to present a full literature review here, but a number of the pertinent papers are discussed and referenced in the following review.

Transformation in Maraging Steels

The phase transformations that are of the most interest for the 18 % Ni maraging steels are the martensite transformation on cooling and the formation of austenite on heating (holding at temperature). As shown in **Figure 1**, martensite is quite stable during heating, which makes possible the aging of the martensite. Data cited for T-250 Maraging steels in **Table 1**, gives A_s and A_f temperature bounds for the $\alpha+\gamma$ region of 661 °C and 730 °C (1223 and 1346 °F), respectively.[1,2,3] However, substantial amounts of reverted austenite can form in Co-free maraging steels (and other maraging steels) during aging treatments at temperatures of less than the A_s temperature.

Table 1: Transformation temperature (reported in Sarma paper, from references)

Steel	M_s C(F)	M_f C (F)	A_s C (F)	A_f C (F)
T-250 Co-free	253 (487)	115 (239)	662 (1223)	730 (1346)
M-250 7.8 Co	209 (410)	90 (194)	630 (1166)	720 (1328)

There are several ways to introduce austenite into the microstructure of martensitic steels: (1) isothermal heating in the two phase austenite + ferrite region, where austenite nucleates and grows (at Ni-rich precipitates and lath boundaries), and (2) thermal cycling at predetermined heating and cooling rates between the single phase austenite region and room temperature, which results in enriched austenite that does not transform to martensite on cooling. In the first case, the austenite is referred to as reverted austenite, in the second as retained austenite. The nickel-enriched austenite has been reported to be stable down to −415 °F (77 K), and is suspected of having high chemical inhomogeneity. In addition, austenite appears to be hardened by high dislocation densities that result from phase work hardening (due to an apparent α-martensite → γ cooperative transformation mechanism that has some similarities to a martensitic transformation).

It is not clear whether reverted or retained austenite adversely affect the scatter in CVN energy. Some authors report no effect of the stable austenite on the impact toughness, but some others report beneficial effects. The effects are likely to be unlike those for retained austenite in other carbon and alloy steels.

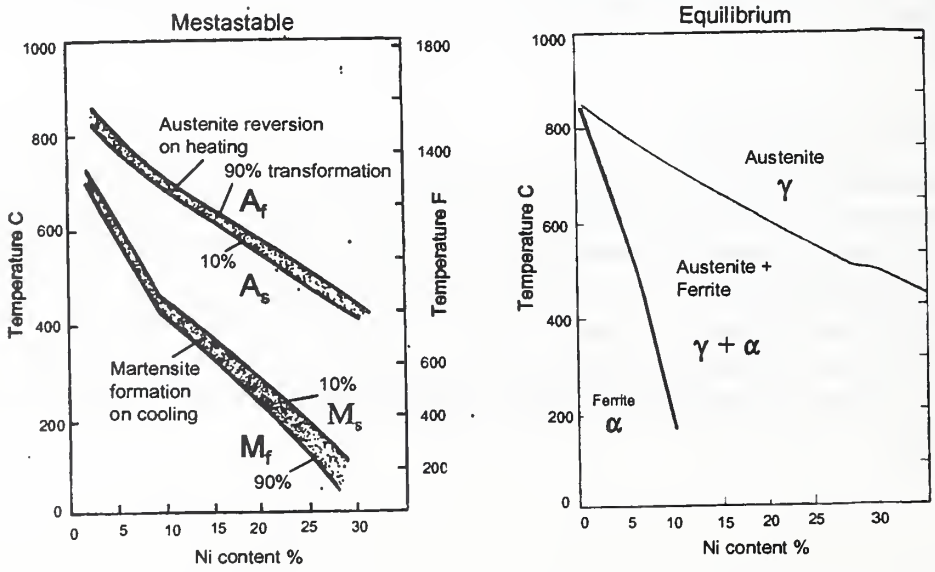


Figure 1. Metastable and equilibrium phase relationship in the Fe –Ni system. These diagrams are based on those shown for maraging steels in the Metals Handbook, Desk Edition, (ASM Metals Park, Ohio), p. 4-57.

Studies of temperature cycling by Viswanathan show that maraging steels are sensitive to rates of heating and cooling.[4] This study used a steel that was solution-annealed at 950 °C (1750 °F) for two hours, air-cooled, then annealed at 820 °C (1500 °F) for 3.5 hours and air-cooled. The steel was then either conventionally aged at 510 °C (950 °F) for 3 hours or thermally cycled and water quenched. Thermal cycling was done at 6 °C per min or 9.5 °C per min or 11.5 °C per min. Cycling was between room temperature and A_f (determined to be 1400 °F; at a heating rate of 750 °C). The retained austenite was found to increase with the number of cycles. The faster cycling rate produced more retained austenite (66 % after five cycles, but at 9.5 °C/min, a single cycle produced about 40 %. The cycling produces solute-rich austenite that does not transform on cooling and this results in a less saturated martensitic phase, which reduces the precipitate strengthening. The Charpy V-notch toughness increased from 12 to 70 J (0 to 60 % Aus) for specimens in the aged condition as the amount of austenite went from 0 to 60 %.

Grain size and morphology

Work by Sinha on a Co-free (250 grade) maraging steel showed the effect of grain size on toughness.[5] The steel had a composition of 0.008 C, 17.1 Ni, 2.25 Mo, 1.39 Ti, 0.01 Al, 0.01 S, 0.008 P, 0.004 O_2 , and 0.003 N_2 . In the study, hot-rolled pieces were solution-annealed at 7 different temperatures for 1 h (air-cooled) and evaluated for microstructure, strength, impact toughness and fracture toughness. Some of the specimens were aged at 477 °C (890 °F) and some were tested in the unaged condition. Full recrystallization occurred after holding for an hour at 825 °C (1520 °F), and this treatment resulted in the optimum strength/toughness combination (full ductile dimple rupture, no ridges, 25 µm blocky martensite). Grain growth occurred at temperatures above 852 °C (1565 °F), accompanied by a gradual change in the martensite lath morphology from blocky to stringer type. The transformation corresponded to a grain size of 35-40 µm. The transformation to stringer type was complete at 1052 °C (1925 °F). Interestingly, this transformation was correlated to a reduction in tensile ductility, fracture toughness, and CVN. There was another decrease in toughness for the 1052 °C (1925 °F) specimens that was attributed to precipitation at grain boundaries. The CVN energy was the least changed of the toughness indicators measured, and was most constant for specimens annealed between the temperatures of 850 and 1000 °C (1565 and 1835 °F). All the CVN data, however, were for aged specimens. (It was also determined in the Sinha study that grain size has little effect on strength, because the martensite lath spacing was not changed by change in grain size.)

Another study by Sinha (T-250 Co-free and M-250 7.5 Co) detailed grain-growth behavior for maraging steel.[6] In this study, isothermal annealing temperatures were used: the specimens were initially annealed for 1 h at 825 °C (1520 °F) and air cooled, then held for times of 0.25 to 10 h at temperatures ranging from 900 to 1050 °C (1655 to 1925 °F). The results show that only modest grain growth (less than 50 µm) occurred for the T-250 Co-free steel at 900 °C (1655 °F) for times up to 3 h. Longer holding times, even at these low temperatures, were shown to sometimes result in abnormal grain growth.

It seems to be generally recognized that grain refinements can be attained in maraging steels by cyclic heating and cooling treatments. Specimens with large grain sizes (hundreds of

micrometers) can be refined to some minimum size (less than 50 micrometers), after which cyclic treatments do not result in further refinement. The paper by Saul addresses this issue for the 250 and 300 grade maraging steels.[1] However, no paper we reviewed presented clear evidence of how and why the refinement occurs, or why the particular treatment schedule is apparently so dependent on the specific composition (or initial microstructure) of the steel.

Aging

Most research does not include aging data for temperatures as low as those used to produce NIST T-200 impact verification specimens (315° C), because aging temperatures this low are not of commercial interest. There has been some indication, however, that aging at low temperatures results in the formation precipitates different from those typical of the 480 °C (900 °F) aging treatment that is commonly used for these alloys.

Studies on an 18 Ni Co-containing 350 grade maraging steel and on T-250 Co-free maraging steel showed differences in the precipitates formed above and below 450 °C (845 °F).[7,8] The studies indicate that Ni_3Ti precipitates are formed in the alloys at high aging temperatures, but at low aging temperatures (315 °C, 3 h) actual precipitation probably does not occur. It is more likely that clusters of Ni and Ti atoms cause the strengthening. The study by Sinha, which include aging temperatures as low as 468 °C on T-250 Co-free maraging steel, shows that the hardening due to aging was rapid (increases of 80 to 90 % within the first 15 to 30 min). Interestingly, the toughness of the under-aged maraging steels in the Sinha study was lower than the toughness for the peak-aged steels. This is apparently due to the clusters or coherent precipitates that are present in the under-aged condition, which restrict cross slip in the matrix. In the peak-aged condition, precipitates (Ni_3Ti) are formed that allow more homogeneous slip in the matrix.

Thermal embrittlement

Maraging steel can become embrittled during high-temperature solution-annealing treatments. The embrittlement is caused by precipitation of Ti (C, N) at grain boundaries during cooling, and can be retained even following re-annealing. Quenching from high temperature prevents the precipitation and subsequent embrittlement.

Sinha studied thermal embrittlement in a T-250 maraging steel and showed that marked degradation in toughness can result when the steel is cooled from high temperature and held between 785 and 400 °C (1450 and 1750 °F).[9] In the study, two heat treatments were used: (1) HT1, solution-treated at 1200 °C (2192 °F) for 1 h and quenched to intermediate temperature for a hold of 180 min then air-cooled, (2) HT2, solution-treated at 1200 °C (2192 °F) and water quenched, then reheated to intermediate temperature for 180 min and air cool. The composition of the T250 steel used in the study was 0.006 C, 0.005 P, 0.001 S, 2.25 Mo, 17.1 Ni, 0.10 Al, 0.003 N, 1.39 Ti, and 0.004 O. The steels were tested in unaged and aged conditions (age at 480 °C for 3 h). The impact energy of the HT1 as-quenched steel was 188 J, compared to 25 J for the HT1 specimens that were held at intermediate temperatures and embrittled. No effect of embrittlement was found at any intermediate temperature for the HT2 treatment (about 190 J for

all intermediate treatments). So embrittlement occurred only when the steel was directly cooled to an intermediate holding temperature.

Other embrittlement studies have shown that the steel must be quenched from re-annealing treatments to avoid embrittlement, but Sinha reasons that in his study there was uniform precipitation of Ti (C,N) on the dislocations formed during the transformation to austenite on the reheating of the solution-annealed and quenched steel (and this kept the Ti out of solution, where it could not segregate to grain boundaries during cooling from the re-annealing treatment).

Inclusions

One would hope that a solution-treatment could be used, in concert with controlled heating and cooling to dissolve and redistribute the large Ti(N,C) inclusions in the T-type maraging steels. This matter is of practical interest because these inclusions can have a significant effect on the homogeneity of the initiation and propagation of ductile tearing in the steel. The solution treatment could also help to redistribute chemical inhomogeneity in the material that might reduce variation in impact properties.

Summary of Past Heat Treatments on the NIST T-200 Bar Stock

A number of heat treatments have been done on the T-200 material at NIST and by NIST contractors. Results from these heat treatments have contributed to our general understanding of this particular heat of T-200 steel, and we discuss some specific details below.

Initial heat treatments on the T-200 material provided a general understanding of the energy levels that might be expected. The mechanical test results for two of the heat treatments are shown in **Figures 2-4.** In **Figures 2 and 3**, the specimens were annealed at 955 °C (1750 °F) for 1 h and air-cooled, then re-annealed at 760 °C (1400 °F) for 1 h and air cool. These specimens were then divided into five groups and aged at 260 °C (500 °F), 288 °C (550 °F), 315 °C (600 °F), 343 °C (650 °F), and 370 °C (700 °F) for 3 h and air-cooled. The data show the relationship between the impact energy and the hardness of the specimens, and indicate that specimens aged at less than 300 °C can reach toughness levels near 200 J.

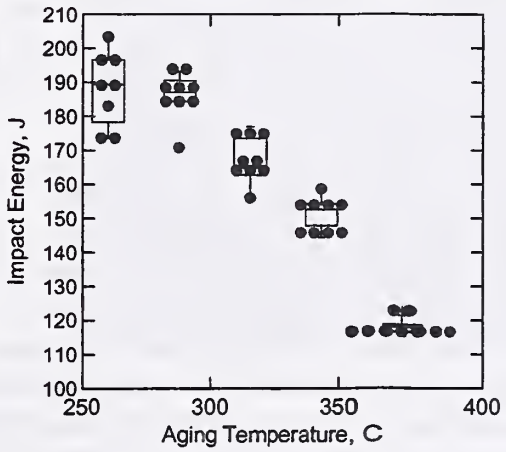


Figure 2: Impact energy versus aging temperature.

The data in **Figure 4** are similar to that in **Figure 2**, but these specimens were annealed at 900 °C (1650 °F) for 1 h and water-quenched, then reheated twice to 675 °C (1250 °F) and water-quenched as a grain refinement treatment, and re-annealed at 815 °C (1500 °F) for 1 h and air-cooled prior to aging at 315 °C (600 °F) and then at 370 °C (700 °F) for 3 h. Other variations

of these two heat-treatment schedules produced similar results. Overall, it appears that this T-200 material can be annealed and aged to produce Charpy specimens having impact energies of 200 J. The hardness of specimens with impact energies of 115 to 205 J ranged from 40 to 32 HRC.

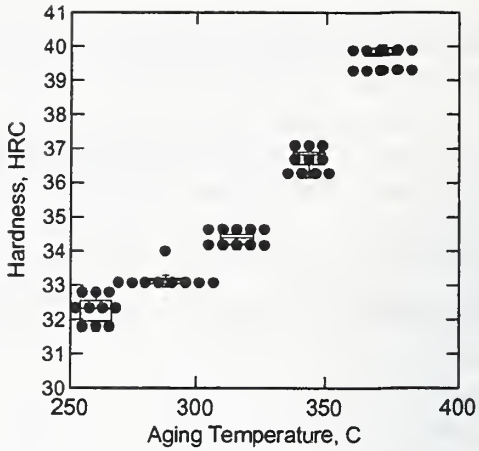


Figure 3: Hardness distributions.

As shown in **Figure 5**, quenching from the annealing temperature clearly results in a softer (tougher) material, and the difference in hardness is retained after aging (compared with air-cooled). These data are from specimens heat-treated at NIST in laboratory furnaces. The specimens were produced mainly for microstructure evaluations, but hardness tests were made to give some indication of the toughness. We found a difference of about 5 HRC, between the quenched and unquenched specimens, which may be helpful in increasing the toughness of the lots. Based on the results shown in **Figures 2 and 3**, an increase in toughness on the order of 30 or 40 J might be associated with a difference in hardness of 5 HRC.

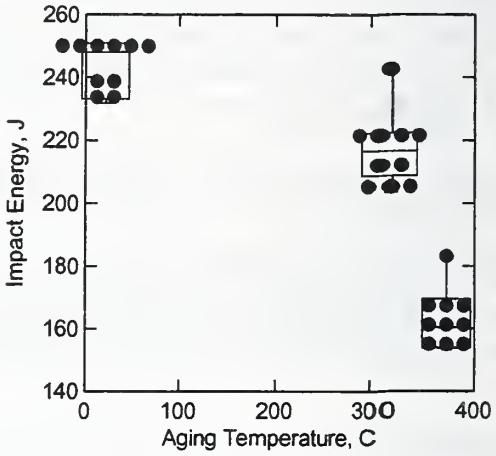


Figure 4: Absorbed energy distributions.

Microstructural evaluations on laboratory specimens suggested limits on annealing temperatures to control grain growth, treatments to refine the grain size, and procedures to control retained austenite levels in the specimens. Examples of the microstructures observed for the specimens are given in **Figure 6-13**.

The as-received T-200 bar stock (**Figure 6**) has a small grain size (likely 10 µm or less), which is desirable. The grain boundaries are decorated with particles or retained austenite. Based on microstructural observations of the specimens heat-treated in our laboratory furnace, grain refinement is attainable. For example, the microstructure shown in **Figure 7**, which has a grain size of about 25 µm or so, was produced from a microstructure having an initial grain size of about 50 µm. In this case, the grain refinement was attained by slowly cooling the specimen through the two-phase region (to room temperature), then reheating and holding it at the temperature where reverted austenite forms, prior to re-annealing the specimen.

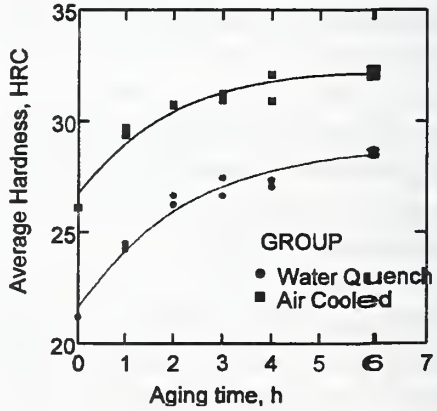


Figure 5: Average hardness trends.

Figure 6. Microstructure of the as-received T-200 bar stock. Bar equal to 10 µm.

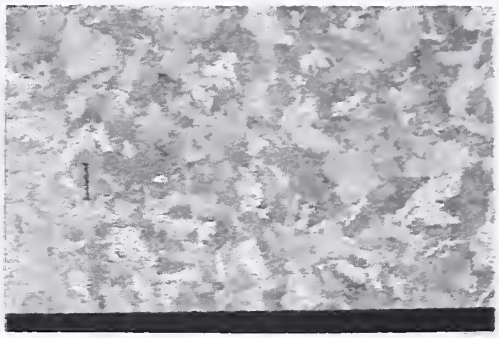


Figure 7. Specimen was annealed at 870 °C air-cooled, re-heated to 575 °C and air cooled , then re-annealed at 840 °C and air-cooled. Bar is equal to 20 µm.

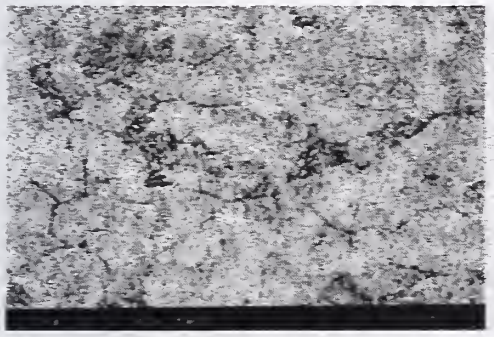


Figure 8. Specimen was annealed at 870 °C and air-cooled, then re-heated to 660 °C for 1 h and water-quenched. Bar equal to 20 µm.

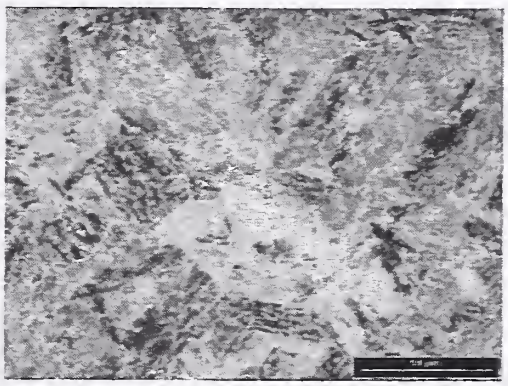


Figure 9. Specimen was annealed at 890 °C and water quenched, then re-heated to 590 °C for 1 h and water-quenched. Bar equal to 50 µm.

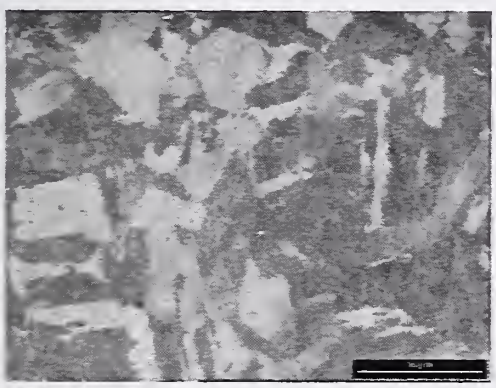


Figure 10. Specimen annealed at 890 °C and water-quenched. Bar equal to 50 µm.

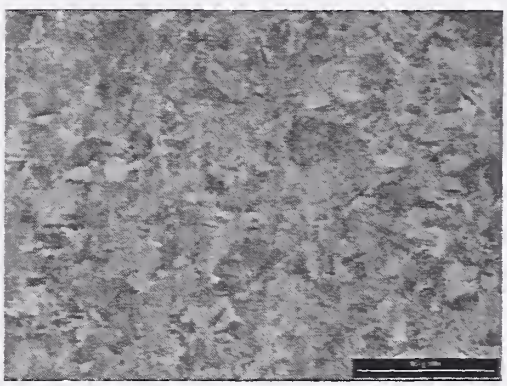


Figure 11. Specimen was heated to 600 °C for 1 h, water-quenched, then reheated to 660 °C for 30 min, then to 840 °C and held 1 h and water-quenched. Bar equal to 50 µm.

In **Figures 8 and 9**, results of holding the specimens at temperatures below the A_f temperature are shown. In **Figure 8**, the microstructure of the specimen, which was held in the two-phase $\alpha + \gamma$ region, has clearly delineated austenite grain boundaries. The basic morphology of the grains has changed to an equiaxed austenite-like morphology. However, the grains contain a fine two-phase structure of reverted austenite and α-martensite, and occasional regions of a coarser two-phase structure. In **Figure 9**, the microstructure of the specimen, which was held at a temperature just below the two-phase region, has significant amounts of reverted austenite at prior austenite grain boundaries and between laths of martensite within the grains.

We did not determine the A_s or the A_f temperature in our experiments, but the estimates given by Sarma in **Table 1** appear to be reasonable for our T-200 material. Our observations also indicate that specimens containing reverted austenite could not be fully annealed when treated at 760 °C (1400 °F) with holding times of 1 h. So, we might consider 815 °C (1500 °F) to be a minimum temperature for annealing treatments. A maximum annealing temperature of about 870 °C (1600 °F) is suggested by our testing, because we see significant grain growth for annealing treatments done at 925 °C (1700 °F) for 1 h, and some grain growth likely also occurred for the 890 °C (1650 °F) annealing treatments, as indicated by the microstructure shown in **Figure 10.**

Overall, the laboratory heat-treatment experiments showed that temperature cycling between room temperature and the two-phase region (and below A_s) yielded some grain refinement in our T-200 material. However, cycles between room temperature to slightly above A_f also yielded some grain refinement, and this cycle avoids the formation of too much reverted austenite. We found that when grain growth occurred due to annealing or solution treatments, it was possible with additional heat treatments to refine the grain size with addition heat treatments back to a reasonably small size (25 µm).

The final topic of discussion here is the grain morphology, because there are two characteristic grain morphologies typical of the specimens. Grain morphologies with little grain-boundary definition (like shown in **Figure 10**) and a resolvable stringer-type martensite (not shown) occur in specimens annealed at higher temperatures. The literature indicates that the morphology is a function of grain size, where the morphology changes at grain sizes near 35 or 40 µm. At smaller grain sizes, the specimens tend to have grain morphologies more like those shown in **Figure 11**. Here, the grain boundaries were etched to mark and delineate the grain boundaries, and the martensite morphology is blocky (according to Sinha) which likely results in the different appearance of the grains. Our T-200 verification specimens that have had the lowest variations in impact energy have had this small grain morphology with well delineated grain boundaries.

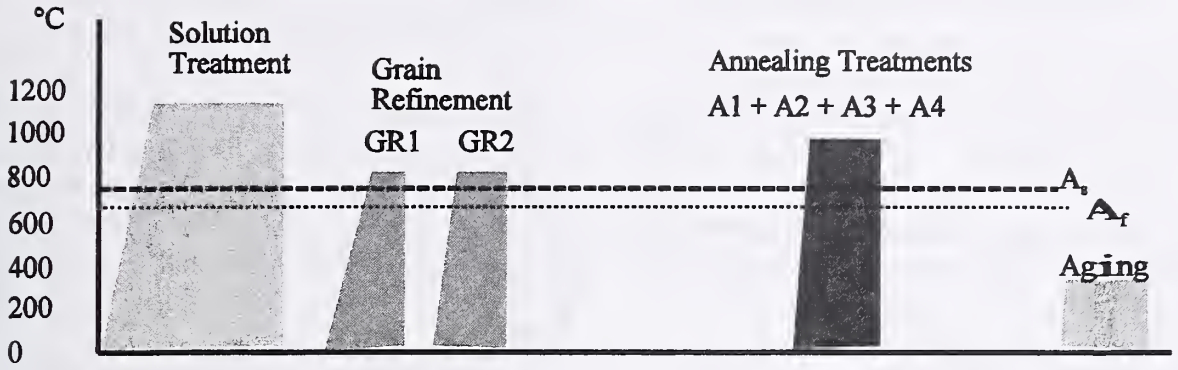


Figure 12: Heat treatment sequence used with reference to the A_s and A_f temperatures.

Materials and Heat Treatments

Three groups of Charpy V-notch specimens were heat-treated by a commercial shop for this study. There were approximately 70 specimens in each group. As shown in **Figure 12** and **Table 2**, the initial heat treatment for the group 1 specimens was as follows: (1) solution treated at 1204 °C (2200 °F) for 1 h and cooled using a 10 bar helium quench, and (2) grain-refinement treatment consisting of a short anneal (GR1) at 815 °C (1500 °F) with a 10 min hold and a 10 bar helium quench, followed by a second anneal (GR2) at 815 °C (1500 °F) for 30 minutes with a 10 bar helium quench. The group 2 specimens did not receive the solution treatment, but did receive the grain refinement treatments (GR1 and GR2). The group 3 specimens received neither the solution treatment nor the grain refinement treatments. So the test matrix has three main legs: (1) group 1, which received a full solution treatment and then grain-refinement steps prior to re-annealing, (2) group 2, which received the same grain refinement steps as the G1 specimens but no solution treatment, and (3) group 3, which was a simple annealing schedule using the as-received material.

Table 2. Heat treatments for the three groups of 1, 2, and 3 specimens. The five specimens taken for evaluation and plotted in Figure 13 are identified.

Group1	Group2	Group3
1204° C ,1 h 10 bar He Specimen 1		
GR1 treatment, 815° C , 10 min, 10 bar He Specimen 2	GR1 treatment, 815° C , 10 min, 10 bar He Specimen 4	
GR2 treatment, 815° C , 30 min, 10 bar He Specimen 3	GR2 treatment, 815° C , 30 min, 10 bar He Specimen 5	
Final anneals (4 variations)	Final anneals (4 variations)	Final anneals (4 variations)
Aging	Aging	Aging

Subgroups of groups 1, 2, and 3, containing 15 to 20 specimens each, were annealed together in the same basket. The four annealing practices were as follows: (1) 830 °C (1525 °F) for 2 h

with a 5 bar helium quench, (2) 830 °C (1525 °F) for 2 h with a 1 bar nitrogen quench, (3) 900 °C (1650 °F) for 2 h with a 5 bar helium quench, and (4) 900 °C (1650 °F) for 2 h with a 1 bar nitrogen quench.

The effect of aging time and temperature on the impact toughness of the specimens was not evaluated. All of the specimens from groups 1, 2, and 3 were aged together at 600 °F for 3 h and quenched in nitrogen at 6 bar.

Specimens for mechanical testing and microstructural evaluations were removed from the group 1 and 2 specimens prior to annealing and aging of the specimens. One specimen was removed from the group 1 specimens following the solution treatment. Another group 1 specimen was removed following the GR1 grain-refinement treatment and still another following the GR2 grain-refinement step. Similarly, two group 2 specimens were removed following the GR1 and GR2 grain-refinement steps, respectively.

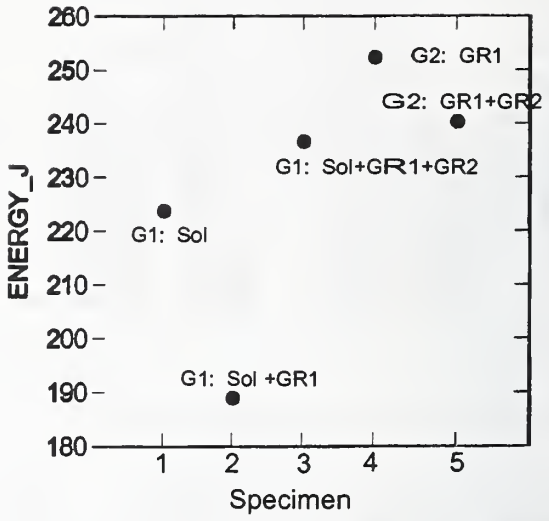


Figure 13. Specimens tested at various stages of the heat treatment are identified by specimen numbers assigned in Table 2. The group (G1 or G2), solution treatment (sol), and grain refinement steps (GR1 and GR2) for the specimens are also indicated.

Results and Discussion

Solution Treatments and Grain Refinement

The results for the initial heat treatments of the group 1, 2, and 3 specimens indicate that the solution treatment of the T-200 did not result in significantly increased toughness for the material. As shown in **Tables 1 and 2**, and in **Figure 13**, the solution-treated G1 specimen, and the solution-treated and grain-refined G1 specimens had respective impact energies of 224 and 237 J. The G2 specimen, which was not solution treated, had an absorbed energy of 240 J following the grain-refinement treatment. These results are from single specimens, but comparing the solution-treated and grain-refined specimen to the as-received and grain-refined specimen, little difference in the level of toughness is apparent. This implies that the as-received material is a relatively homogenous bulk material and solution treatments might be expected to have only limited effect on the toughness level of the specimens.

The microstructure of the group 1 specimen (**Figure 14**) that was evaluated following the solution treatment at 1204 °C (2200 °F) had a very large grain size, as might be expected. The microstructure appears blocky rather than a stringer type, however, which may indicate that grain size does not always dictate the grain morphology. It was also noted that the grain boundaries

etch unevenly, and this may indicate that precipitates or slight chemical inhomogeneities are present at the boundaries even for the hardest quenched specimens.

The group 1 specimen that was solution treated, then re-heated slowly to 815 °C (1500 °F) for 10 min had a predominantly stringy martensitic structure (**Figure 15**). The group 1 specimen (**Figures 16 and 17**) that was solution annealed and re-heated twice for grain refinement had a mixture of blocky and stringer-type structures. More importantly, the grain size is clearly refined compared to that for the as-solution-treated structure (**Figure 14**). This specimen had the highest toughness.

The group 2 specimens (**Figures 18-19**), which were not solution treated, show slight differences in structure from one another. The specimen that received only one grain-refinement step, **Figure 18**, has reverted austenite at prior austenite grain boundaries and on preferred planes within the grains (as does the as-received T-200 bar stock). This specimen had the highest energy for the heat-treatment conditions considered here. The specimen that received both grain-refinement treatments (**Figure 19**) has a slightly larger grain size, and less reverted austenite is apparent. Both specimens retained a reasonably small grain size (probably less than 25 µm). The slightly higher toughness of the specimen in **Figure 18**, may reflect the smaller grain size and/or the presence of more reverted austenite in the structure.

Comparing the group 1 and 2 specimens, grain-refinement treatments following the solution-treatment were effective in reducing the very large grain size of the as-solution-treated specimens. The large grain in the center of the solution treated specimen shown in **Figure 14** is about 500 µm in diameter, compared with grain diameters on the order of 10 to 20 µm in the nonsolution-treated group 2 specimens (**Figures 18 - 19**). It is not clear whether any grain refinement occurred in the group 2 specimens. These specimens have grain sizes similar to the as-received T-200 material.

The solution treatment did not result in a change to the large indigenous Ti(N,C) inclusion content or size distribution, as might be expected (the melting point of these inclusions is over 2900 °C). Measurements on group 1 and group 3 specimens (400 fields per sample) show the Ti-rich inclusions to favor cube-like morphologies with an average size of about 11 µm (cube edge). The average number of Ti(C,N) inclusions per millimeter squared was estimated to be about 10. There were differences in the amounts of smaller indigenous inclusions in the samples. The as-received sample had a significantly higher number of inclusions with diameters in the range of about 2 to 30 µm. Detailed evaluations of these smaller inclusions (which may also include small islands of retained austenite) were not done, however, so no data on these inclusions are available (the inclusion counts were done at too low a magnification to yield accurate information on these smaller inclusions).

Since the solution treatment does not result in a beneficial modification to the large Ti-rich inclusions, and the as-received material was not found to be embrittled, it appears that there is little need to include a solution treatment in the processing of this material, unless it results in lower scatter to the specimens following the final annealing treatments.

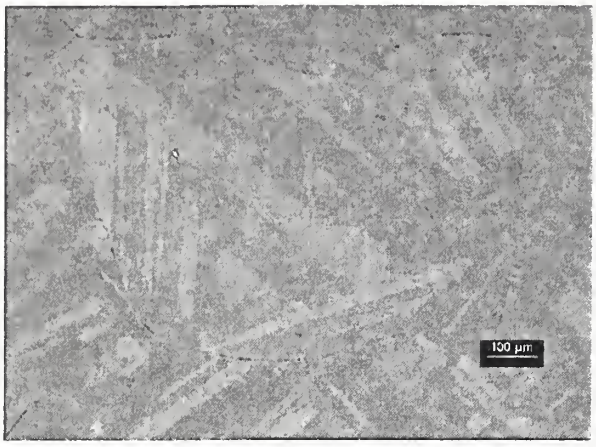


Figure 14. Specimen G1-1

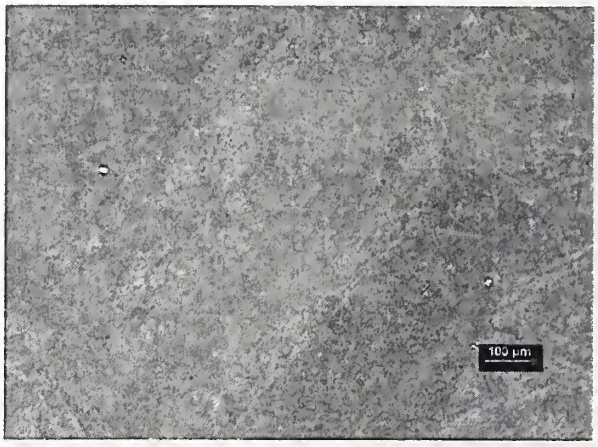


Figure 15. Specimen G1-2

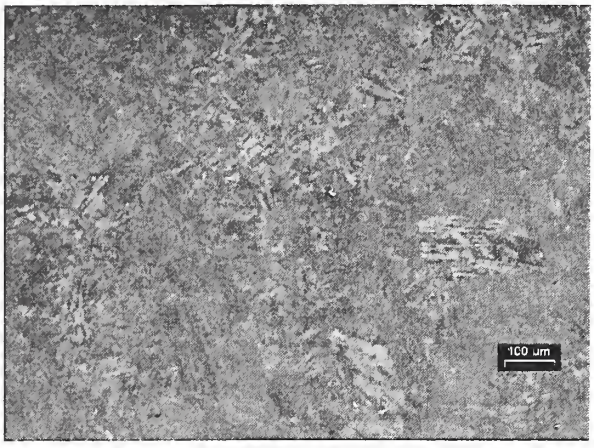


Figure 16. Specimen G1-3

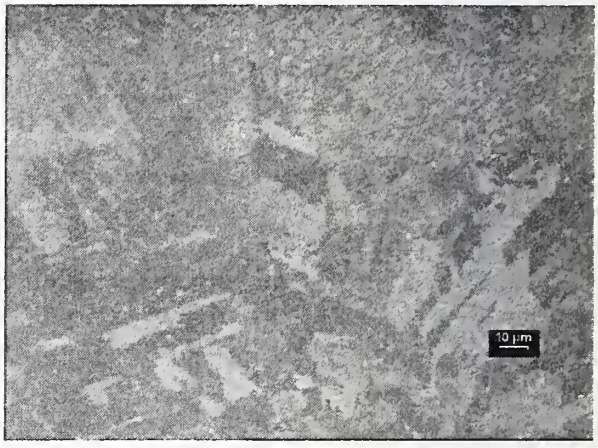


Figure 17. Specimen G1-3

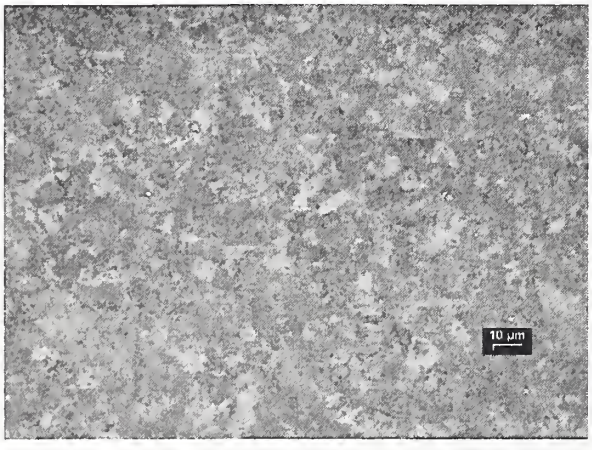


Figure 18. Specimen G2-1

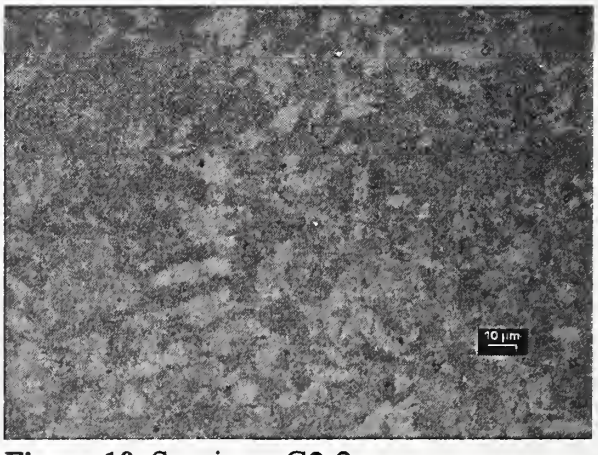


Figure 19. Specimen G2-2

Final Annealing **treatments (1 through 4)**

The microstuctures of the specimens after the final annealing treatments showed that the group 2 and group 3 specimens had smaller grain size than the group 1 specimens. In **Figure 20**, for example, the grain size of the group 1 specimen is larger than that of the group 3 specimens shown in **Figures 21 and 22**. The group 1 specimens, the only specimen group that was solution-annealed, had some grains as large as 100 μm, and many grain diameters were assumed to be between 20 and 40 μm.

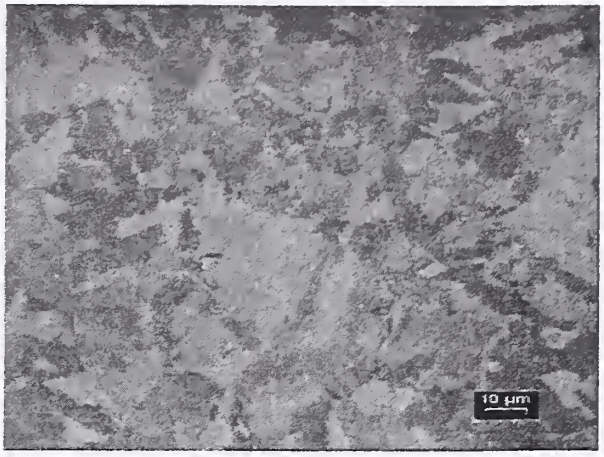


Figure 20. A group 1 specimen, with a high-temperature anneal, specimen # 391. Bar equal to 10 μm.

The group 2 and group 3 specimens had similar grain sizes. However, those specimens that were annealed at the lower temperature (**Figure 21**) generally had a smaller grain size (typical sizes range between 4 and 20 μm in **Figure 21**) than those annealed at higher temperature (typical sizes range between 15 to 30 μm). So, some grain growth was associated with the final annealing treatments, particularly for the higher-temperature anneals (900 °C, 1650 °F), but grain sizes remained reasonable for all of the treatments.

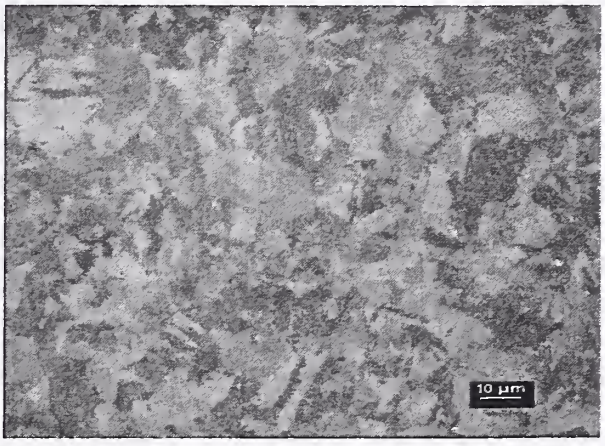


Figure 21. A group 3 specimen, with a low-temperature anneal, specimen # 2.

The impact-test results are summarized in **Figure 23.** The results of anneal 1 (1525 °F, 2 h, 5 bar helium quench) showed that a slightly higher absorbed energy (157 ft-lbs) was attained for the group 1 specimens, which were solution-annealed. The group 3 specimens, however, had lower scatter in absorbed energy than the group 1 or group 2 specimens. So the particular solution and grain refinement treatments used here for the Group 1 and 2 specimens did not reduce the scatter in impact energy over that found for the as-received and annealed specimens.

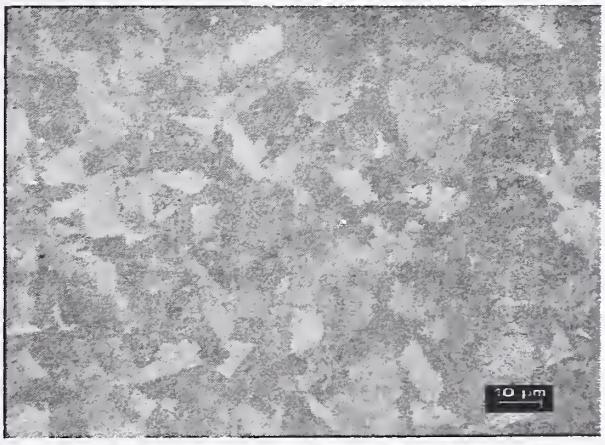


Figure 22. A group 3 specimen with high-temperature anneal (#36)

The results of anneal 2 (1525 °F, 2 h, 1 bar nitrogen quench) again show that the group 1 specimens have slightly higher toughness (153 ft-lbs) than the group 2 or group 3 specimens (147 and 149 ft-lbs). However, the slower quench used for this annealing step resulted in higher scatter

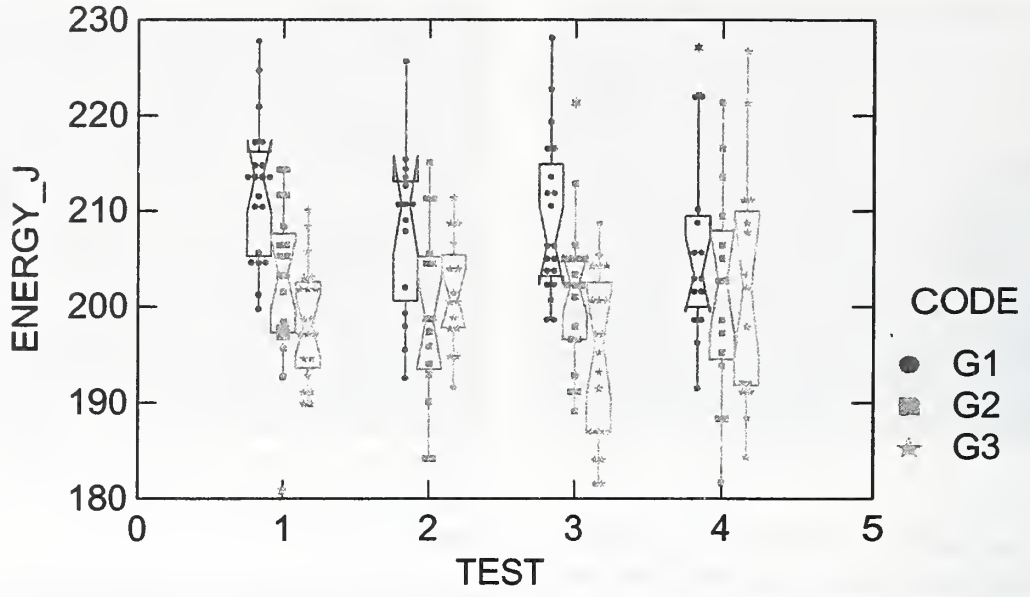


Figure 23. Box plot of the data for the final annealing treatments.

(coefficient of variation) for groups 1 and 2 (0.042 and 0.047), compared with the results for anneal 1 (0.035 and 0.040). The scatter for the group 3 specimens is the exception here, where similar low levels of scatter (0.03) were found to be independent of cooling rate.

The anneal 3 (1650 °F, 2 h, 5 bar helium quench) produced data of impact energy similar to that for the anneal 1 and 2 treatments: 154, 149, and 144 ft-lbs for groups 1, 2, and 3, respectively. The scatter in absorbed energy for the group 1 and 2 specimens (0.039 and 0.039) was similar to previous results, but the scatter for the group 3 specimens increased (0.045).

The data for the anneal 4 specimens (1650 °F, 2 h, 1 bar nitrogen quench) showed absorbed energy levels similar to all the previous data, but had consistently higher scatter in absorbed energy: The coefficients of variation for group 1, 2, and 3 were 0.049, 0.055, and 0.062, respectively. Again, the slower quench rate (used to simulate an air cool for this annealing treatment) resulted in higher scatter for the specimens.

Summary

Results of this study and the initial studies on this T-200 material have provided useful information for the production and quality control of the super-high-energy specimens. A summary of our understanding for the new T-200 material is as follows:

- The T-200 material is relatively homogeneous.
- The T-200 material can be used to produce impact verification specimen having energies of near 200 J.
- A minimum temperature 815 ° C (1500 °F) is suggested for annealing treatments.
- A maximum annealing temperature of about 870 ° C (1600 °F) is suggested.
- Significant grain growth occurs at temperatures above 900 °C (1700 °F, for 1 h).

- Grain refinement is possible if grain growth occurs during heat treatment.
- A small, more or less equiaxed grain morphology (less than 30 µm) with well defined grain boundaries is desirable.
- Increasing the amount of reverted austenite in the microstructure appears to increase the toughness of the material (but the effect on variation in the absorbed energy was not evaluated)
- The variation in the absorbed energy is likely reduced by quenching the material rather than allowing it to cool more slowly.
- The variation in the absorbed energy was not clearly reduced by solution treatments or by grain refinement treatments.

CODE	NUM	ANNEAL	Anneal code	ENERGY (ft-lbf)	Energy (J)	HARD (HRC)
G1	1	SOL	0	165	224	25
G1	2	SOL_GR1	0	139	189.0	30
G1	3	SOL_GR1_GR2	0	175	237	28
G2	1	GR1_ONLY	0	186	252	28
G2	2	GR1_GR2	0	177	240	27
G1	4	GR1_5BAR	1	158	214	30
G1	5	GR1_5BAR	1	158	214	30
G1	6	GR1_5BAR	1	158	215	30
G1	7	GR1_5BAR	1	147	200	30
G1	8	GR1_5BAR	1	160	217	30
G1	9	GR1_5BAR	1	151	205	30
G1	10	GR1_5BAR	1	155	211	30
G1	11	GR1_5BAR	1	166	225	30
G1	12	GR1_5BAR	1	158	214	30
G1	13	GR1_5BAR	1	152	206	30
G1	14	GR1_5BAR	1	160	218	
G1	15	GR1_5BAR	1	148	201	
G1	16	GR1_5BAR	1	156	212	
G1	17	GR1_5BAR	1	151	204	
G1	18	GR1_5BAR	1	168	228	
G1	19	GR1_5BAR	1	151	205	
G1	20	GR1_5BAR	1	159	215	
G1	21	GR1_5BAR	1	155	210	
G1	22	GR1_5BAR	1	163	221	
G1	23	GR1_5BAR	1	158	214	
G2	3	GR1_5BAR	1	152	206	31
G2	4	GR1_5BAR	1	152	206	31
G2	5	GR1_5BAR	1	146	198	31
G2	6	GR1_5BAR	1	133	181	31
G2	7	GR1_5BAR	1	146	198	31
G2	8	GR1_5BAR	1	151	205	31
G2	9	GR1_5BAR	1	150	203	31
G2	10	GR1_5BAR	1	156	212	31
G2	11	GR1_5BAR	1	149	202	31
G2	12	GR1_5BAR	1	146	198	31
G2	13	GR1_5BAR	1	153	207	
G2	14	GR1_5BAR	1	145	197	
G2	15	GR1_5BAR	1	158	214	
G2	16	GR1_5BAR	1	144	196	
G2	17	GR1_5BAR	1	156	211	
G2	18	GR1_5BAR	1	154	208	
G2	19	GR1_5BAR	1	142	193	
G2	20	GR1_5BAR	1	145	197	
G2	21	GR1_5BAR	1	158	214	
G2	22	GR1_5BAR	1	150	203	
G3	1	GR1_5BAR	1	140	190	31
G3	2	GR1_5BAR	1	143	194	31
G3	3	GR1_5BAR	1	141	191	31
G3	4	GR1_5BAR	1	142	193	31
G3	5	GR1_5BAR	1	150	203	31
G3	6	GR1_5BAR	1	146	198	31
G3	7	GR1_5BAR	1	149	202	31
G3	8	GR1_5BAR	1	141	191	31

CODE	NUM	ANNEAL	Anneal code	ENERGY (ft-lbf)	Energy (J)	HARD (HRC)
G3	9	GR1_5BAR	1	150	203	31
G3	10	GR1_5BAR	1	154	208	31
G3	11	GR1_5BAR	1	149	202	
G3	12	GR1_5BAR	1	140	190	
G3	13	GR1_5BAR	1	155	210	
G3	14	GR1_5BAR	1	149	202	
G3	15	GR1_5BAR	1	147	199	
G3	16	GR1_5BAR	1	152	206	
G3	17	GR1_5BAR	1	146	198	
G3	18	GR1_5BAR	1	145	197	
G3	19	GR1_5BAR	1	144	195	
G3	20	GR1_5BAR	1	145	197	
G1	24	GR1_1BAR	2	144	195	30
G1	25	GR1_1BAR	2	149	202	30
G1	26	GR1_1BAR	2	155	211	30
G1	27	GR1_1BAR	2	153	208	30
G1	28	GR1_1BAR	2	142	193	30
G1	29	GR1_1BAR	2	157	213	31
G1	30	GR1_1BAR	2	159	215	30
G1	31	GR1_1BAR	2	166	226	30
G1	32	GR1_1BAR	2	155	211	30
G1	33	GR1_1BAR	2	156	211	30
G1	34	GR1_1BAR	2	158	214	30
G1	35	GR1_1BAR	2	158	214	30
G1	36	GR1_1BAR	2	146	198	30
G1	37	GR1_1BAR	2	154	209	30
G1	38	GR1_1BAR	2	147	199	30
G2	23	GR1_1BAR	2	156	211	31
G2	24	GR1_1BAR	2	159	215	31
G2	25	GR1_1BAR	2	144	196	31
G2	26	GR1_1BAR	2	143	194	31
G2	27	GR1_1BAR	2	142	193	31
G2	28	GR1_1BAR	2	146	197	31
G2	29	GR1_1BAR	2	136	184	31
G2	30	GR1_1BAR	2	146	198	31
G2	31	GR1_1BAR	2	140	190	31
G2	32	GR1_1BAR	2	156	212	31
G2	33	GR1_1BAR	2	152	206	
G2	34	GR1_1BAR	2	151	204	
G2	35	GR1_1BAR	2	147	199	
G2	36	GR1_1BAR	2	136	184	
G2	37	GR1_1BAR	2	151	205	
G3	21	GR1_1BAR	2	148	200	31
G3	22	GR1_1BAR	2	151	204	31
G3	23	GR1_1BAR	2	146	198	31
G3	24	GR1_1BAR	2	156	211	31
G3	25	GR1_1BAR	2	147	199	31
G3	26	GR1_1BAR	2	152	207	31
G3	27	GR1_1BAR	2	150	204	31
G3	28	GR1_1BAR	2	144	195	31
G3	29	GR1_1BAR	2	148	201	31
G3	30	GR1_1BAR	2	141	192	31
G3	31	GR1_1BAR	2	144	195	

CODE	NUM	ANNEAL	Anneal code	ENERGY (ft-lbf)	Energy (J)	HARD (HRC)
G3	32	GR1_1BAR	2	154	209	
G3	33	GR1_1BAR	2	149	201	
G3	34	GR1_1BAR	2	154	208	
G3	35	GR1_1BAR	2	146	198	
G1	39	GR2_5BAR	3	160	217	30
G1	40	GR2_5BAR	3	168	228	30
G1	41	GR2_5BAR	3	147	199	30
G1	42	GR2_5BAR	3	150	204	30
G1	43	GR2_5BAR	3	148	201	29
G1	44	GR2_5BAR	3	156	212	30
G1	45	GR2_5BAR	3	153	207	30
G1	46	GR2_5BAR	3	151	204	30
G1	47	GR2_5BAR	3	155	210	30
G1	48	GR2_5BAR	3	147	199	30
G1	49	GR2_5BAR	3	152	206	
G1	50	GR2_5BAR	3	150	203	
G1	51	GR2_5BAR	3	164	223	
G1	52	GR2_5BAR	3	149	202	
G1	53	GR2_5BAR	3	151	205	
G1	54	GR2_5BAR	3	162	219	
G1	55	GR2_5BAR	3	160	216	
G1	56	GR2_5BAR	3	157	212	
G1	57	GR2_5BAR	3	158	214	
G1	58	GR2_5BAR	3	152	205	
G2	38	GR2_5BAR	3	140	189	30
G2	39	GR2_5BAR	3	141	192	31
G2	40	GR2_5BAR	3	151	205	31
G2	41	GR2_5BAR	3	148	201	30
G2	42	GR2_5BAR	3	151	205	31
G2	43	GR2_5BAR	3	151	205	31
G2	44	GR2_5BAR	3	145	197	31
G2	45	GR2_5BAR	3	150	203	31
G2	46	GR2_5BAR	3	149	202	30
G2	47	GR2_5BAR	3	145	197	31
G2	48	GR2_5BAR	3	157	213	
G2	49	GR2_5BAR	3	146	198	
G2	50	GR2_5BAR	3	152	205	
G2	51	GR2_5BAR	3	152	206	
G2	52	GR2_5BAR	3	141	191	
G2	53	GR2_5BAR	3	163	221	
G2	54	GR2_5BAR	3		0	
G2	55	GR2_5BAR	3	149	202	
G2	56	GR2_5BAR	3	142	193	
G2	57	GR2_5BAR	3	150	203	
G3	36	GR2_5BAR	3	151	204	31
G3	37	GR2_5BAR	3	138	187	31
G3	38	GR2_5BAR	3	138	187	31
G3	39	GR2_5BAR	3	145	197	31
G3	40	GR2_5BAR	3	136	184	31
G3	41	GR2_5BAR	3	151	204	31
G3	42	GR2_5BAR	3	138	187	31
G3	43	GR2_5BAR	3	144	195	31
G3	44	GR2_5BAR	3	141	192	31

CODE	NUM	ANNEAL	Anneal code	ENERGY (ft-lbf)	Energy (J)	HARD (HRC)
G3	45	GR2_5BAR	3	151	205	31
G3	46	GR2_5BAR	3	134	182	
G3	47	GR2_5BAR	3	148	201	
G3	48	GR2_5BAR	3	154	209	
G3	49	GR2_5BAR	3	148	201	
G3	50	GR2_5BAR	3	136	184	
G3	51	GR2_5BAR	3	146	198	
G3	52	GR2_5BAR	3	143	193	
G3	53	GR2_5BAR	3	134	181	
G3	54	GR2_5BAR	3	152	205	
G3	55	GR2_5BAR	3	148	201	
G1	59	GR2_1BAR	4	149	201	30
G1	60	GR2_1BAR	4	149	202	30
G1	61	GR2_1BAR	4	147	199	30
G1	62	GR2_1BAR	4	152	205	30
G1	63	GR2_1BAR	4	147	199	30
G1	64	GR2_1BAR	4	152	206	30
G1	65	GR2_1BAR	4	150	203	30
G1	66	GR2_1BAR	4	150	203	30
G1	67	GR2_1BAR	4	141	192	30
G1	68	GR2_1BAR	4	154	209	30
G1	69	GR2_1BAR	4	168	227	
G1	70	GR2_1BAR	4	164	222	
G1	71	GR2_1BAR	4	145	196	
G1	72	GR2_1BAR	4	164	222	
G1	73	GR2_1BAR	4	155	210	
G2	58	GR2_1BAR	4	143	194	
G2	59	GR2_1BAR	4	151	205	
G2	60	GR2_1BAR	4	139	189	
G2	61	GR2_1BAR	4	139	188	
G2	62	GR2_1BAR	4	158	214	
G2	63	GR2_1BAR	4	144	195	
G2	64	GR2_1BAR	4	150	203	
G2	65	GR2_1BAR	4	147	199	
G2	66	GR2_1BAR	4	163	221	
G2	67	GR2_1BAR	4	134	182	
G2	68	GR2_1BAR	4	155	209	
G2	69	GR2_1BAR	4	150	203	
G2	70	GR2_1BAR	4	146	197	
G2	71	GR2_1BAR	4	152	206	
G2	72	GR2_1BAR	4	160	217	
G3	56	GR2_1BAR	4	153	208	
G3	57	GR2_1BAR	4	142	193	
G3	58	GR2_1BAR	4	150	203	
G3	59	GR2_1BAR	4	136	184	
G3	60	GR2_1BAR	4	141	191	
G3	61	GR2_1BAR	4	149	202	
G3	62	GR2_1BAR	4	142	192	
G3	63	GR2_1BAR	4	154	209	
G3	64	GR2_1BAR	4	156	211	
G3	65	GR2_1BAR	4	146	198	
G3	66	GR2_1BAR	4	167	227	
G3	67	GR2_1BAR	4	163	221	

CODE	NUM	ANNEAL	Anneal code	ENERGY (ft-lbf)	Energy (J)	HARD (HRC)
G3	68	GR2_1BAR	4	139	188	
G3	69	GR2_1BAR	4	141	192	
G3	70	GR2_1BAR	4	156	212	

Anneal code	Group CODE	STATISTIC	ENERGY J (ft-lbf)	HARD (HRC)
0	G1	N of cases	1	1
0	G1	Mean	234 (165)	25
0	G1	N of cases	1	1
0	G1	Mean	188 (139)	30
0	G1	N of cases	1	1
0	G1	Mean	237 (175)	28
0	G2	N of cases	1	1
0	G2	Mean	252 (186)	28
0	G2	N of cases	1	1
0	G2	Mean	240 (177)	27
1	G1	N of cases	20	10
1	G1	Minimum	199 (147)	30
1	G1	Maximum	228 (168)	30
1	G1	Range	28 (21)	1
1	G1	Median	214 (158)	30
1	G1	Mean	213 (157)	30
1	G1	Standard Dev	7 (5)	0
1	G1	C.V.	0	0
1	G2	N of cases	20	10
1	G2	Minimum	180 (133)	31
1	G2	Maximum	214 (158)	31
1	G2	Range	34 (25)	0
1	G2	Median	203 (150)	31
1	G2	Mean	202 (149)	31
1	G2	Standard Dev	8 (6)	0
1	G2	C.V.	0	0
1	G3	N of cases	20	10
1	G3	Minimum	190 (140)	31
1	G3	Maximum	210 (155)	31
1	G3	Range	20 (15)	0
1	G3	Median	198 (146)	31
1	G3	Mean	198 (146)	31
1	G3	Standard Dev	5 (4)	0
1	G3	C.V.	0	0
2	G1	N of cases	15	15
2	G1	Minimum	193 (142)	30
2	G1	Maximum	225 (166)	31
2	G1	Range	32 (24)	1
2	G1	Median	210 (155)	30
2	G1	Mean	207 (153)	30
2	G1	Standard Dev	10 (7)	0
2	G1	C.V.	0	0
2	G2	N of cases	15	10
2	G2	Minimum	184 (136)	31
2	G2	Maximum	216 (159)	31
2	G2	Range	32 (23)	1
2	G2	Median	198 (146)	31
2	G2	Mean	199 (147)	31
2	G2	Standard Dev	10 (7)	0
2	G2	C.V.	0	0
2	G3	N of cases	15	10
2	G3	Minimum	191 (141)	31
2	G3	Maximum	212 (156)	31
2	G3	Range	21 (15)	0

Anneal code	Group CODE	STATISTIC	ENERGY J (ft-lbf)	HARD (HRC)
2	G3	Median	201 (148)	31
2	G3	Mean	202 (149)	31
2	G3	Standard Dev	5 (4)	0
2	G3	C.V.	0	0
3	G1	N of cases	20	10
3	G1	Minimum	199 (147)	29
3	G1	Maximum	228 (168)	30
3	G1	Range	29 (22)	1
3	G1	Median	206 (152)	30
3	G1	Mean	209 (154)	30
3	G1	Standard Dev	8 (6)	0
3	G1	C.V.	0	0
3	G2	N of cases	19	10
3	G2	Minimum	190 (140)	30
3	G2	Maximum	221 (163)	31
3	G2	Range	33 (24)	1
3	G2	Median	202 (149)	31
3	G2	Mean	202 (149)	31
3	G2	Standard Dev	8 (6)	0
3	G2	C.V.	0	0
3	G3	N of cases	20	10
3	G3	Minimum	182 (134)	31
3	G3	Maximum	209 (154)	31
3	G3	Range	27 (20)	1
3	G3	Median	197 (145)	31
3	G3	Mean	195 (144)	31
3	G3	Standard Dev	8 (6)	0
3	G3	C.V.	0	0
4	G1	N of cases	15	10
4	G1	Minimum	191 (141)	30
4	G1	Maximum	228 (168)	30
4	G1	Range	35 (26)	0
4	G1	Median	203 (150)	30
4	G1	Mean	206 (152)	30
4	G1	Standard Dev	11 (8)	0
4	G1	C.V.	0	0
4	G2	N of cases	15	0
4	G2	Minimum	182 (134)	
4	G2	Maximum	221 (163)	
4	G2	Range	39 (29)	
4	G2	Median	203 (150)	
4	G2	Mean	202 (149)	
4	G2	Standard Dev	11 (8)	
4	G2	C.V.	0	
4	G3	N of cases	15	0
4	G3	Minimum	184 (136)	
4	G3	Maximum	226 (167)	
4	G3	Range	42 (31)	
4	G3	Median	202 (149)	
4	G3	Mean	202 (149)	
4	G3	Standard Dev	12 (9)	
4	G3	C.V.	0	

1. Saul, G.; Roberson, J.A.; Adair, A.M.; The Effects of Thermal Treatment on the Grain Size and Mechanical Properties of 18 Ni Maraging Steels, Met. Trans., Vol. 1, February 1970, pp. 383-387.

2. Viswanathan, U.K.; Kutty, T.R.G.; Ganguly, C.; Dilatometric Technique for Evaluation of the Kinetics of Solid-State Transformation of Maraging Steel, Met. Trans. Vol. 24, December 1993, pp. 2653-2656.

3. Sarma, D.S.; Sinha, P.P.; Structure-Property Correlations in a Co-free Maraging Steel, Trans. Indian Inst. Met., Vol. 49, No.3, June 1996, pp. 163-169.

4. Viswanathan, U.K.; Kishore, R.; Asundi, M.K.; Effect of Thermal Cycling on the Mechanical Properties of 350-Grade Maraging Steel, Metallurgical and Material Transactions A, Vol. 27A, March 1996, pp. 757-761.

5. Sinha, P.P.; Tharian, T.T.; Sivakumar, D.; Sarma, D.S.; Effects of Solution Temperatures on Microstructure, Tensile, and Fracture Properties of Co-free 18 Ni Maraging Steel, Steel Research, 65, Vol. No. 11, 1994, pp. 494-499.

6. Sinha, P.P.; Sreekumar, K.; Nargarajan, K.V.; Grain Growth in 18 Ni Co-free Maraging Steel, Trans. Indian Inst. Met., Vol. 52, No. 2-3, April-June 1999, pp. 119-124.

7. Tewari, S.; Mazumder, I.S.; Batra, G.K.; Dey, Banerjee, S.; Precipitation in 18 Ni Maraging Steel of Grade 350, Acta mater, 48, 2000, pp. 1187-1200.

8. Sinha, K.; Tharian, T.; Sreekumar, K.; Nargarajan, K.V.; Sarma, D.S.; Effect of Aging on Microstructure and Mechanical Properties of Cobalt Free 18Ni250 Maraging Steel, Materials Science and Technology, January 1998, Vol 14, pp. 1-9.

9. Sinha, D;. Sivakumar, K.T.; Tharian, K.V.; Nargarajan, and D.S. Sarma, Thermal Embrittlement in 18Ni Colbalt Free and 18Ni-8Co-5Mo Maraging Steels, Material Science and Technology, November 1996, Vol 12, pp. 945-954.

C.N. McCowan

Evaluation of 11.1 mm Bar Stock for Service in the Charpy V-notch Program

Introduction

The microstructure of the 11.1 mm (7/16 in) bar stock for the Charpy V-notch (CVN) Program was evaluated to determine whether this stock might be used for making specimens. For comparison purposes, samples of the 12.7 mm (1/2 in) bar we have been using in the program, and samples of the 14.3 mm (9/16 in) bar purchased by the Army were also evaluated.

The main purpose of the evaluation was to determine the depths of decarburization, oxidation, and surface flaws in the bar stock. If the 11.1 mm stock is to be considered for use in the Charpy Program, superficial and microstructural discontinuities should be limited to depths of less than 0.5 mm.

Procedure

Several metallographic specimens were cut from bars of 11.1 mm, 12.7 mm, and 14.3 mm stocks. Samples were prepared for evaluation by grinding and polishing. Samples evaluated by optical microscopy were etched with picric acid. Samples evaluated by scanning electron microscopy (SEM) were not etched. A backscatter detector was used to enhance contrast in the SEM between oxide scales and the steel. The Rockwell hardness was measured on the outside surface of the bars and near the center (core) of the cross-sectioned bars. Vickers microhardness was measured with a 10 g load.

Results

11.1 mm × 11.1 mm Stock: Figure 1

The 11.1 mm cold-rolled bar stock made for NIST has: (1) a zone of internal oxidation typically less than 0.01 mm deep, (2) a nearly totally decarburized zone with a depth of about 0.1 mm, (3) a uniform core structure starting at about 0.5 mm, and (4) a surface hardness of about 93 HRB and a core hardness of about 97 HRB.

The microstructure of the bar near the surface is fine-grain, equiaxed ferrite. The adjacent partially decarburized zone is a mixture of ferrite and spheroidized pearlite. The microstructure in the core of the bar is spheroidized pearlite.

12.7 mm × 12.7 mm Stock: Figure 2

The 12.7 mm bar stock made for NIST typically has: (1) internal oxidation to a depth of 0.02 mm with excursions to depths approaching 0.2 mm, (2) a zone of nearly total decarburization about 0.1 mm deep at the surface, (3) a uniform core structure that starts at a depth of about 2 mm, and (4) a surface hardness near 89 HRB and a core hardness of 91 HRB.

The microstructure of the decarburized zone near the surface of the bar is equiaxed ferrite. The microstructure of the partially decarburized zone has pro-eutectoid ferrite at the grain boundaries with partially spheroidized pearlite within the grains. A band structure containing very fine uniformly spheroidized carbides separates the decarburized zone from the core structure (zone at

approximately 1 to 1.5 mm from the surface of the bar). The core structure is predominately pearlite.

14.3 mm × 14.3 mm Stock: Figures 3 and 4
The 14.3 mm bar stock the Army used has: (1) a zone of internal oxidation, (2) an almost totally decarburized region that is similar to that found in our 12.7 mm bar stock (0.05 to 0.1 mm), (3) a uniform core structure that starts at a depth of about 0.6 to 1 mm, and (4) a surface hardness of about 104 R(b) and a core hardness of 110 R(b).

The microstructure of this bar in the initial decarburized zone has large ferrite grains with what appears to be bainite (or martensite) at the ferrite grain boundaries. Adjacent to this region, a zone (band) with a morphology similar to that of the initial decarburized zone is found on a finer scale. The remaining portion of the decarburized zone is a mixture of spheroidized pearlite, ferrite, and bainite. The core microstructure is spheroidized pearlite and bainite.
The hardness of the grain boundary product is 685 HV (81 HRB) compared with 340 and 500 HV (65 and 75 HRB) for the ferrite and pearlite products, respectively.

Discussion
Internal oxidation often occurs at grain boundaries near the surface of the bars. This results in a network of microcracks that may increase the sensitivity of the bar to quench cracking, particularly microcracks surrounding laps (**Figure 2b**). All the bar stocks evaluated have internal oxidation, but it is most severe in the 12.7 mm bar stock.

The oxidation in the bars made for NIST (11.1 mm and 12.7 mm) is typically on the order of several grains in depth ahead of the scale interface, so the very small grain size at the surface of the 11.1 mm bar stock appears to help limit the depth of microcracking.

The 14.3 mm bar has a larger grain size at the surface than the two other bar sizes evaluated. The large grain size appears to be due to grain growth during annealing of the hot-rolled bar. This coarse-grain region was likely formed by the coarsening of the band of microstructure adjacent to it. The network of microcracks present in the sample indicated that a much smaller grain size was present when the internal oxidation occurred.

The specimens from the 12.7 mm bar stock are machined to a final size of 10 mm × 10 mm, so they should be clear of the non-uniform structures found near the surface of the stock. However, in comparison to the "Army stock," the depth of the nonuniform structure in our 12.7 mm bar is about twice as deep.

Producing CVN specimens from the 11.1 mm cold-rolled bar appears to be feasible: a 10 mm × 10 mm specimen from the core of the bar would be free of surface defects and have a uniform structure.

Recommendations
I recommend that we make 100 CVN specimens from our 11.1 mm bar for evaluation. The heat treatment currently used for the low-energy specimens should be used. The absolute energy value of the batch is not important; the variance in energy is what needs to be evaluated at this time.

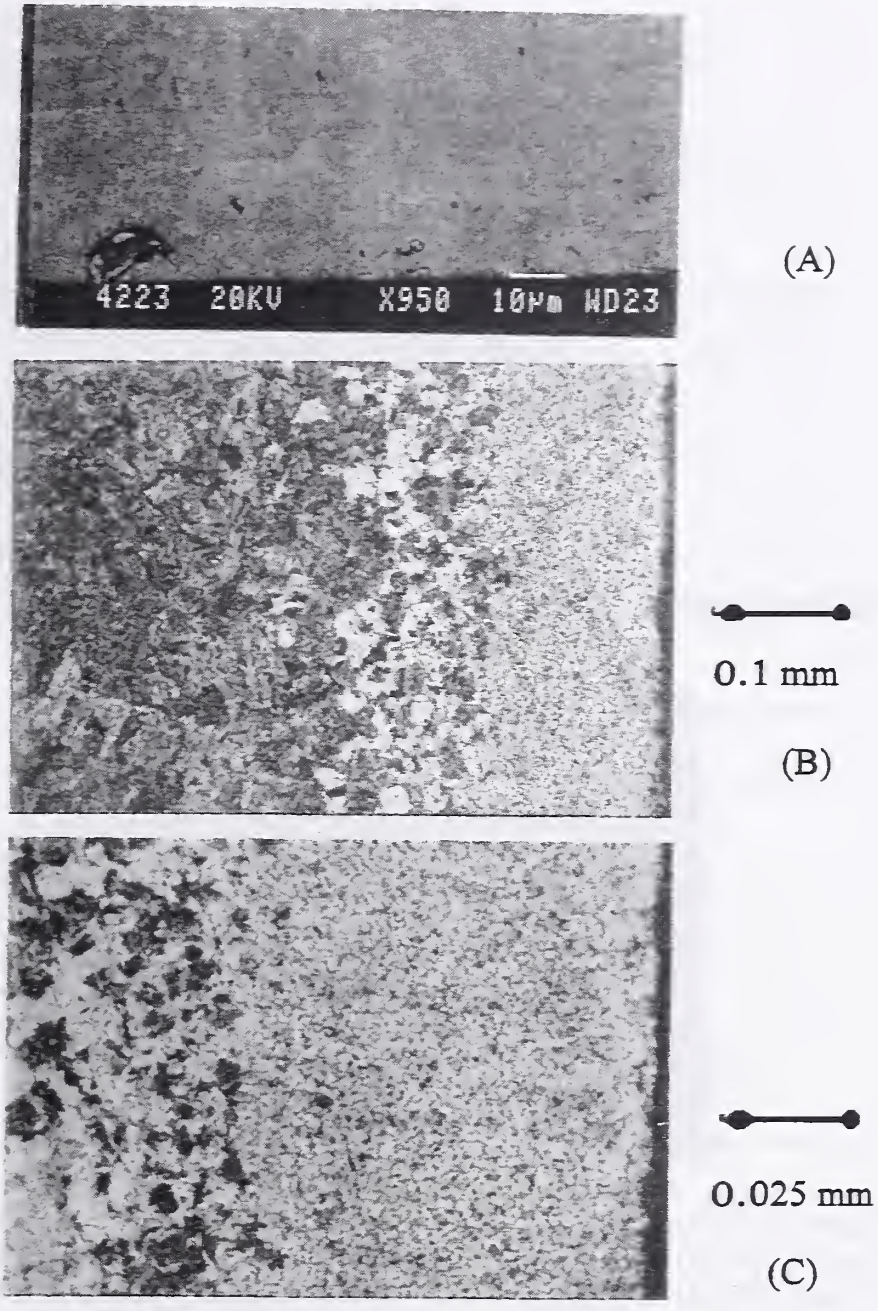


Figure 1. The 11.1 mm bar stock has: (A) shallow microcracking near the surface and surrounding surface defects, (B) a heavily decarburized zone and a partially decarburized zone precending the core structure, and (C) a small equiaxed grain size near the surface of the bar.

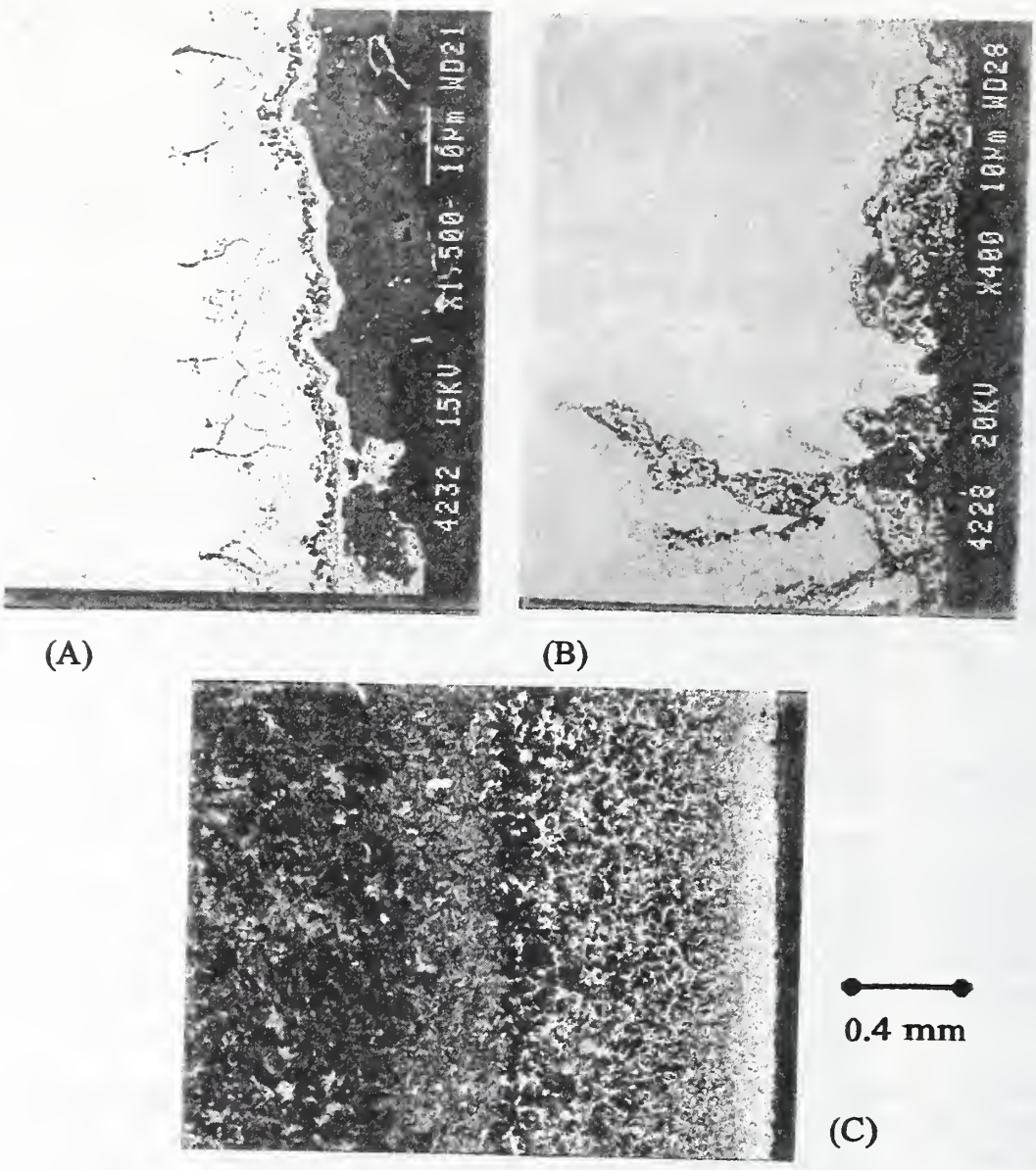


Figure 2. The 12.7 mm bar stock has: (A-B) microcracking due to internal oxidation near surface of the bar, and (C) a heavily decarburized zone followed by a partially decarburized zone, then a band of finely spheroidized carbides preceding the core microstructure.

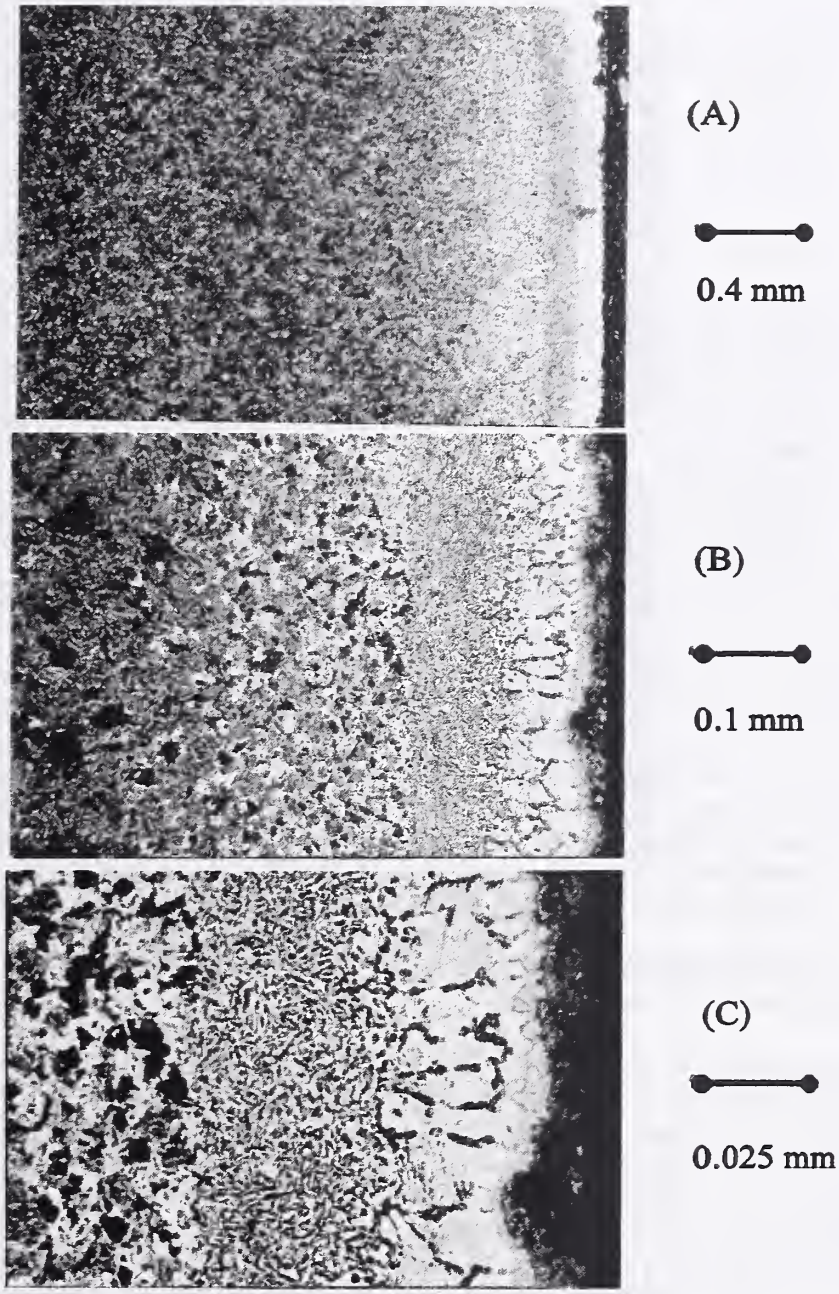


Figure 3. The 14.3 mm bar stock has an elongated-grain morphology near the surface of the bar in the heavily decarburized zone (A). The depth of the micrographing here is shown by the light gray lines just ahead of the oxide scale layer. The darker areas at the elongated grain boundaries are not cracks; these regions are probably a transformation product (bainite or martinsite). Coarse-grain region and the network of microcracks near the surface (B-C).

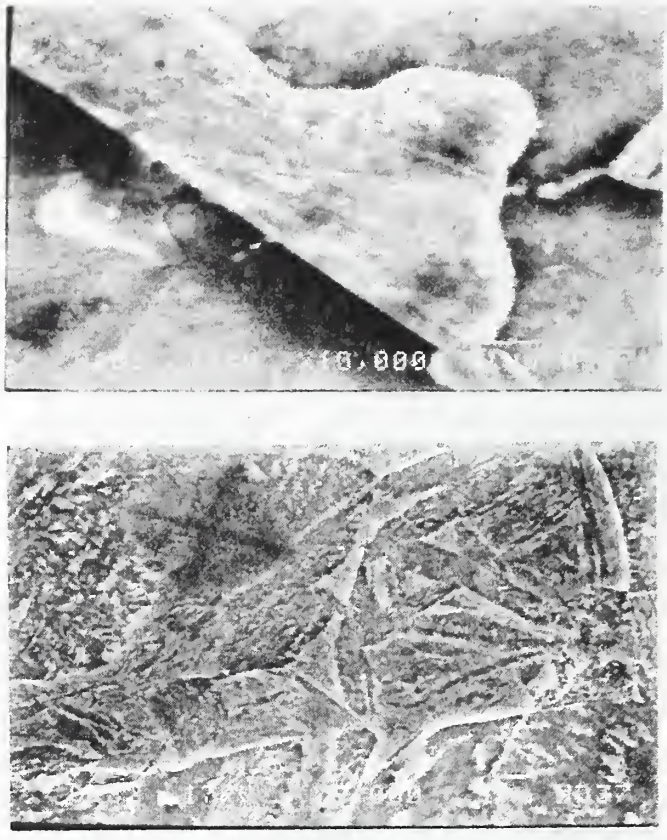


Figure 4. The microstructure of the product present at the grain boundaries near the surface (A) and core (B) of the 14.3 mm bar stock.

George E. Hicho

Investigation of Cracks in Charpy V-notch Specimens

Introduction

On December 18, 1992, Mr. Paul Lundberg of the SRM group here at the Gaithersburg NIST site asked us to help the Boulder NIST group determine the cause of the cracking observed in a number of prospective standard Charpy V-notch specimens prepared from a new heat of 4340 VIM-VAR steel supplied by Vanadium Alloy Steel Corporation of Latrobe, Pennsylvania. After phone conversations with our Boulder Labs, we asked them to send us specimens showing the cracking. Subsequently we received two cracked specimens, one apparently heat-treated but not machined, for it still had surface scale, and one that was heat-treated and machined to finished size. The hardness on both specimens was determined and found to be HRC 25, indicative of quenched and tempered 4340 steel. Macroscopic examinations indicated that quite possibly the cracks could have been present initially, and grew during the quenching procedure. These are the so-called "quench cracks." Another suggestion as to the cause of the cracks was that residual stresses could have been created by the letter-stenciling operation prior to heat treating, and one or a combination of the subsequent heat treatments could have increased the residual stresses in the vicinity of the letters to the point where cracking occurred. And finally there was supposition that decarburization of the specimen's surface could have led to the cracking, and surface cracks could have propagated into the specimen after a heat treatment. Conversations with the Boulder scientists who examined the rods using acoustic apparatus revealed that no deep cracks were present in the as-received steel. With this background, it was decided to metallurgically examine randomly selected stock to determine the as-received microstructure throughout the specimen and to proceed from there.

Examination Procedure

We requested that Dr. Tom Siewert (Charpy project leader) and Dan Vigliotti send us two pieces of the steel rod approximately 15 cm (6 in) in length, from the as-received stock. These rods had no cracks in them. Upon arrival at Gaithersburg, samples were cut from the two bars and prepared for metallographic examination. The samples were sectioned so that we could look into the long axis of the bar that is perpendicular to the rolling direction of the bar. **Figure 1** shows a photomicrograph of the "short bar" observed in this direction. It is evident from the photomicrograph that there was a decarburized layer (white region) encompassing the outer edge of this sample. Decarburization is the removal of elemental carbon from the surface, usually in a heat-treating operation, by exposing the surface to an oxidizing atmosphere. Also evident is the microstructure gradient that started at the decarburized zone and progressed inward toward the center of the bar. This "grey area" (chemical segregation?) was followed by an area that contained what was apparently the as-received annealed microstructure, pearlite plus ferrite. **Figures 2 and 3** show the microstructure at the surface and in the gradient region. Note the presence of the small surface cracks in the decarburized zone in both of these photos. Even though the photomicrographs revealed that the decarburized zone was small, hardness measurements indicated otherwise. Hardness measurements, **Table 1**, taken on the sample indicated that the low-hardness zone was about 2 mm (0.080 in) deep. This indicated that, in addition to the presence of a decarburized layer, quite possibly there was some chemical segregation.

Figure 4 shows a photomicrograph of the "long bar." Some decarburization is present, and hardness measurements revealed that the depth of decarburization was not as deep as that found in the "short bar." **Figures 5 and 6** also show that the decarburized layer of the long bar contained small cracks.

On December 22, we received some heat-treated specimens that had cracks in them. The specimen in the center of the photomicrograph, **figure 7**, is the original specimen which had a decarburized layer (white ring) around the outside of the specimen. It was decided to section the cracked specimens to see whether decarburized and "grey" zones, like those found in the randomly picked bar, were also present in these samples. If so, we could state that the decarburized layer could be causing the cracking seen in the samples. This was not so, for none of the samples surrounding the decarburized sample shown in **figure 7** showed any extensive decarburization. It was then concluded that the decarburized layer did not cause the cracks found in these specimens. The surfaces of the heat-treated and cracked specimens were examined and small surface cracks that did not grow into primary cracks were found. **Figures 8 and 9** show cracks, on these same specimens, and those that did grow after heat treatment.

It was suggested that quite possibly the decarburized and "grey" zones were the result of some chemical inhomogeneity in the steel. Initially, the centers of the specimens that cracked were checked for manganese, chromium, nickel, silicon, and molybdenum contents and compared to the composition of AISI 4340 steel. The composition was determined in each using EDAX. The results indicated that Mn, Cr, Ni, and Mo were within the composition limits for 4340 steel, but the silicon contents were not. The effects of increased silicon on cracking susceptibility is not known at this time.

A chemical profile for Cr, Ni, and Mn from the surface layer containing the decarburized zone through the "grey" zone was performed. The results indicated that there was no significant change in these elements as the zones were traversed. It appears that these zones were depleted primarily of carbon.

Conclusions

Metallographic examinations, and chemical and hardness determinations were conducted on two steel samples that were randomly selected from a batch of rods that were to be made into standard Charpy V-notch specimens. Microstructure and hardness examinations revealed that one bar was severely decarburized. Whether this decarburized layer was the initiator of the crack in the specimens, as some thought, was determined by metallographically examining these as-received specimens for decarburization. The photos showed that there were small surface cracks present in the surface of the as-received and heat-treated specimens. Hardness results showed that the sample's decarburized layer had a lower hardness compared to that of its core since the decarburized layer is essentially ferrite with the carbon removed, while the core is essentially a mixture of ferrite and pearlite. It is our opinion that it is rather difficult for a pre-existing crack to propagate into and through the tough ferrite. There is no fracture mechanics work, to our knowledge, that would indicate that cracks, even if present in the surface, would propagate from the surface into the main body of the specimen. Our initial metallographic examinations on the two cracked samples showed no deep surface cracks, with the exception of what we call "scale cracks." We surmise now that quite possibly, during one of the heat treating processes, one of these small "scale cracks" grew into the crack we saw in the finished specimen.

A second set of cracked specimens was received and metallographically examined. The specimens were cross-sectioned and compared to the specimen that was not cracked, but showed the decarburized and gradient zones (**figure 7**). The cracked specimens did not show the same microstructure near the surface; hence it was concluded that the microstructure was not conducive to promoting the cracking seen in the specimens. The photomicrographs also showed that there were small surface cracks present.

There should be a caveat here, however, that specimens with this amount of decarburized and gradient zones would appear to be unwanted as standard specimens, since their microstructures are not homogeneous. The carburized and gradient zones present in the bars suggests that the bars could have been annealed several times; or, quite possibly during the annealing process, the bars were exposed to an oxidizing atmosphere.

It appears that there were two synergistic causes for the cracking in the specimens. One is that the as-received bars contained small surface cracks, and the second is that stamping the identification code on these specimens induces a residual stress in the specimen that is severely increased in the heat-treating operation. The combination of the small surface cracks and the stresses enhanced during the transformation to untempered martensite appear to cause the cracking in the specimens.

Recommendations

Prior to receiving the heat-treated cracked samples, two questions were to be answered. One was "Does the excessively deep decarburization layer induce cracking in the specimens?" The second was "Does stenciling of the numbers on the sample prior to heat treating initiate cracks?" The first question has been answered: cracks do not initiate in the decarburized layer since we found that none of the heat-treated cracked samples contained an extensive decarburized layer. To resolve the second question, we recommend that the as-received bars that do not contain excessive decarburization, as determined by metallographic examination, be evaluated in a rigid heat-treating experiment. After metallographic examination, each bar free of this decarburization should be cut to Charpy-specimen length, but not machined to size. Five of these ten samples should then be stamped in the usual manner, and the five remaining unstamped samples be sent to the heat treater and included with a batch of samples being heat-treated. In addition to these bars, five specimens should have their surfaces machined and then stamped. These specimens should be added to the others and heat-treated. The samples should be normalized in the usual manner, and then examined carefully for cracks. If none are visible, perform the next recommended hardening heat treatment that is, austenitizing followed by quenching in oil. After this treatment, remove the samples and examine them carefully for cracks. If none is visible, proceed with the sub-zero treatment, remove the samples, and see whether cracking has occurred. Finally, let these 10 samples remain as they are, that is, do not temper them. After an hour, observe which group, if any, are cracked. This experiment should resolve which heat-treatment step is producing the stresses necessary to cause cracking.

This experiment still does not resolve the problem of the gross decarburization found in randomly selected specimens. This problem should be addressed since these specimens were intended to be made into standard Charpy V-notch impact specimens.

Table 1. Tukon hardness results. A 1 kg load and 20x objective were used.

Dist. from Surface (inch)	Long Bar	Remark	Short Bar	Remark
0.003	157	Decarb.	134	Decarb.
0.005	160	"	158	"
0.010	169	"	181	w/ferrite
0.020	187	w/ferrite	198	"
0.040	194	"	214	"
0.050	---	----	219	hazy
0.060	198	w/ferrite	213	hazy
0.080	197	core	198	core
0.100	206	"	208	"
0.140	202	"	202	"
0.180	206	"	205	"
0.220	204	"	210	"

Table 2. Chemical composition (wt %) of 4340 steel. Only selected elements were determined.

Element	AISI	"short bar"	"long bar"
Manganese	0.60/0.80	0.80	0.81
Chromium	0.70/0.90	0.80	0.87
Nickel	1.65/2.00	1.88	1.95
Silicon	0.20/0.35	0.40	0.58
Molybdenum	0.20/0.30	0.18	0.23

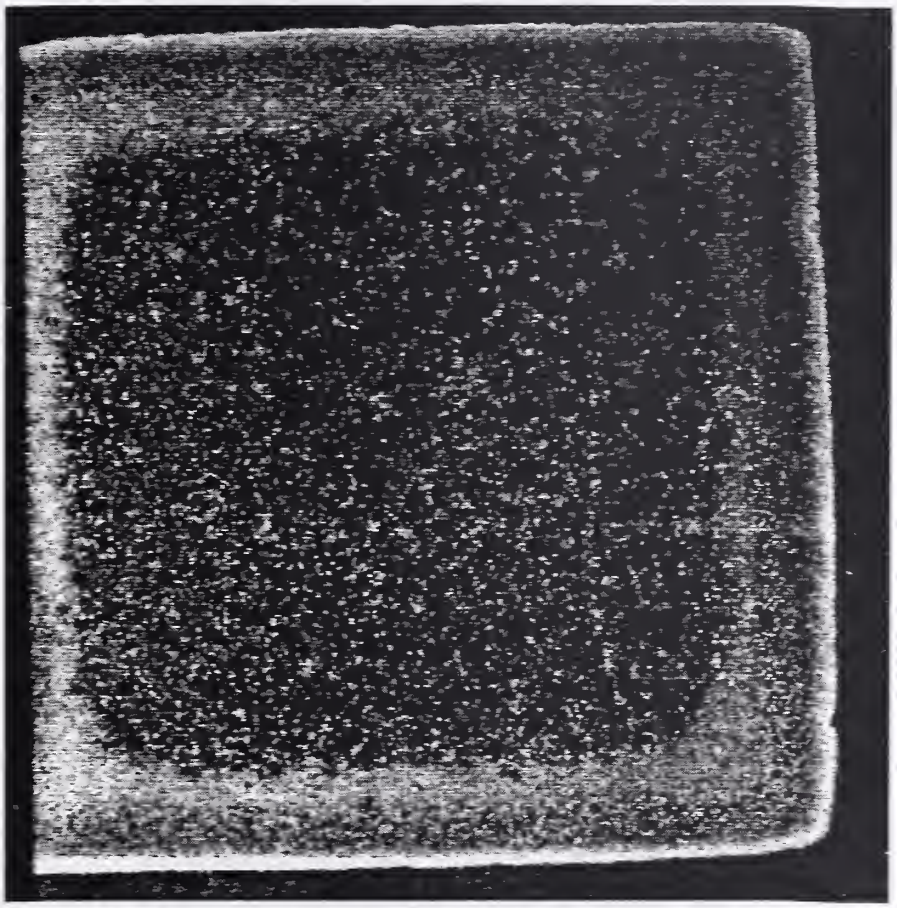


Figure 1. Photomicrograph of the "short bar" showing the outside decarburized layer, the transition region, which could be created by chemical segregation, and the core, consisting of ferrite plus pearlite. Mag. 6.5 X; Etch: 2 % Nital

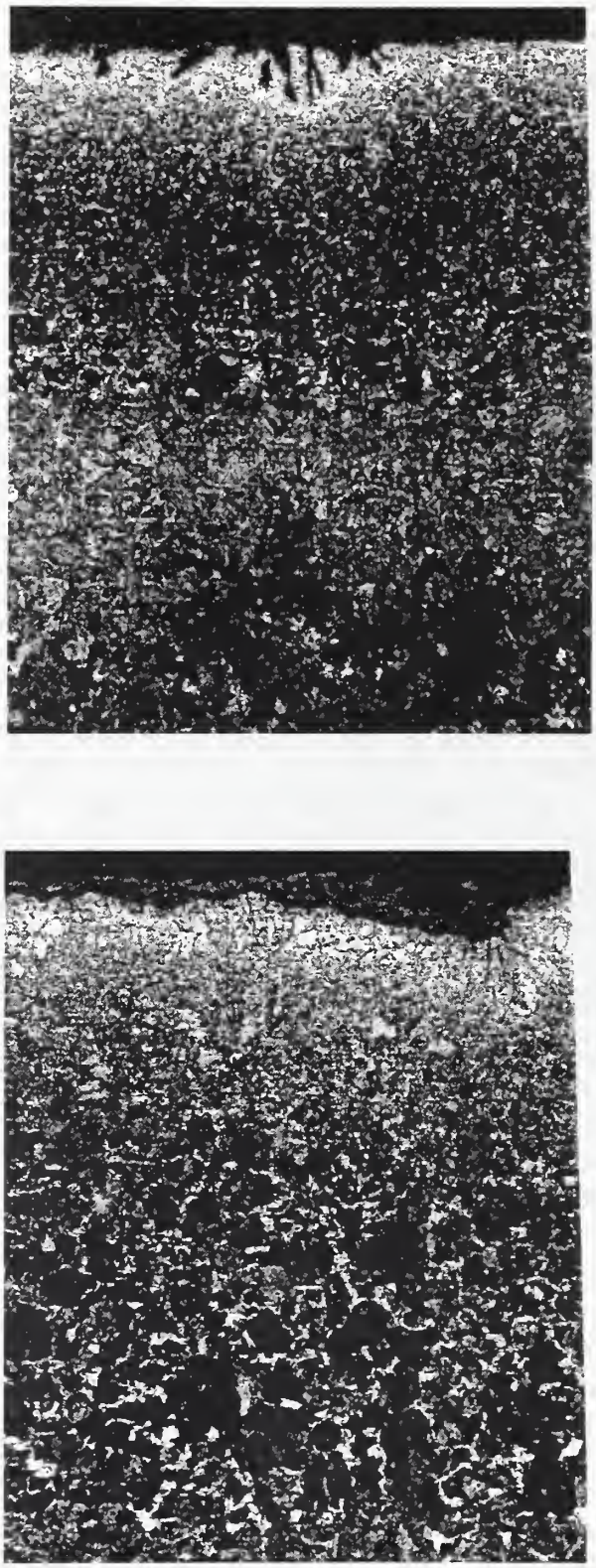


Figure 2. Photomicrographs of the "short bar" showing the decarburized layer and the adjacent microstructure.
Mag: (A) 50 X; (b) 100 X, Etch: 2 % Nital

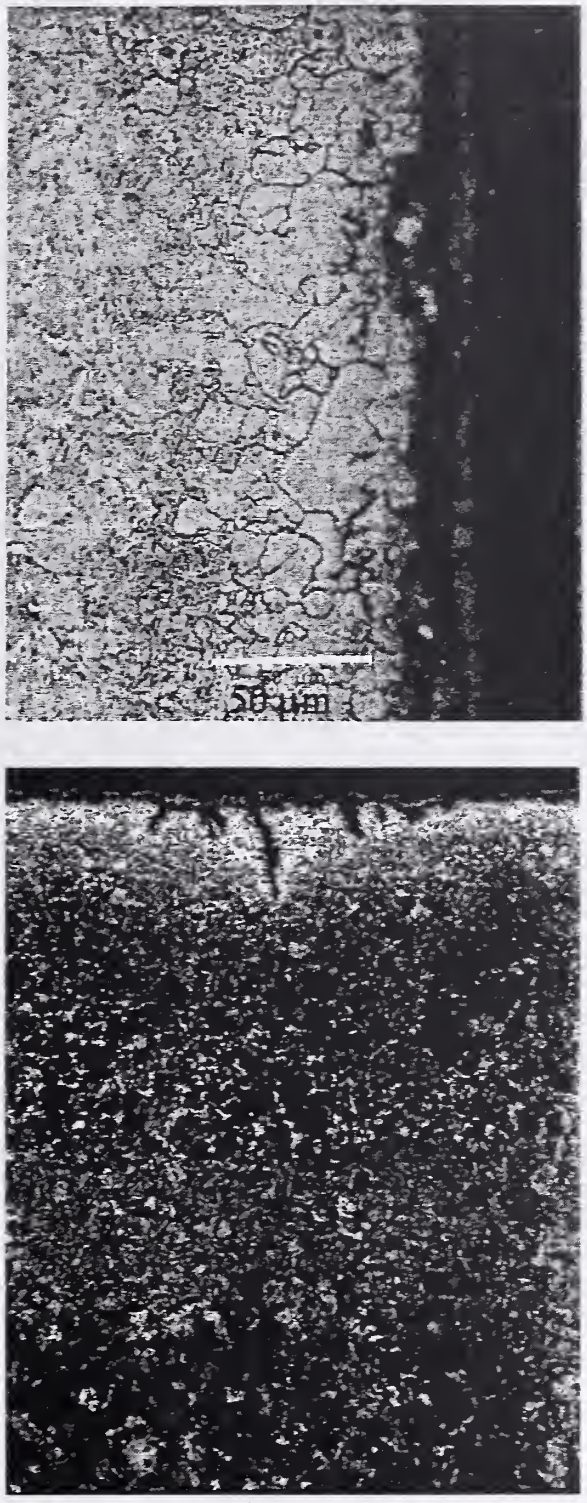


Figure 3. Photomicrographs of the "short bar" at higher magnification. Scale and decarb areas near the surface are shown. Note variation in microstructure from surface to center. Mag: (a) 50 µm bar; (b) 50 X; Etch: 2% Nital

Figure 4. Photomicrograph of "long bar" showing the absence of the extensive decarb layer and transition zone as shown in the "short bar." Mag: 6.5 X; Etch: 2 % Nital

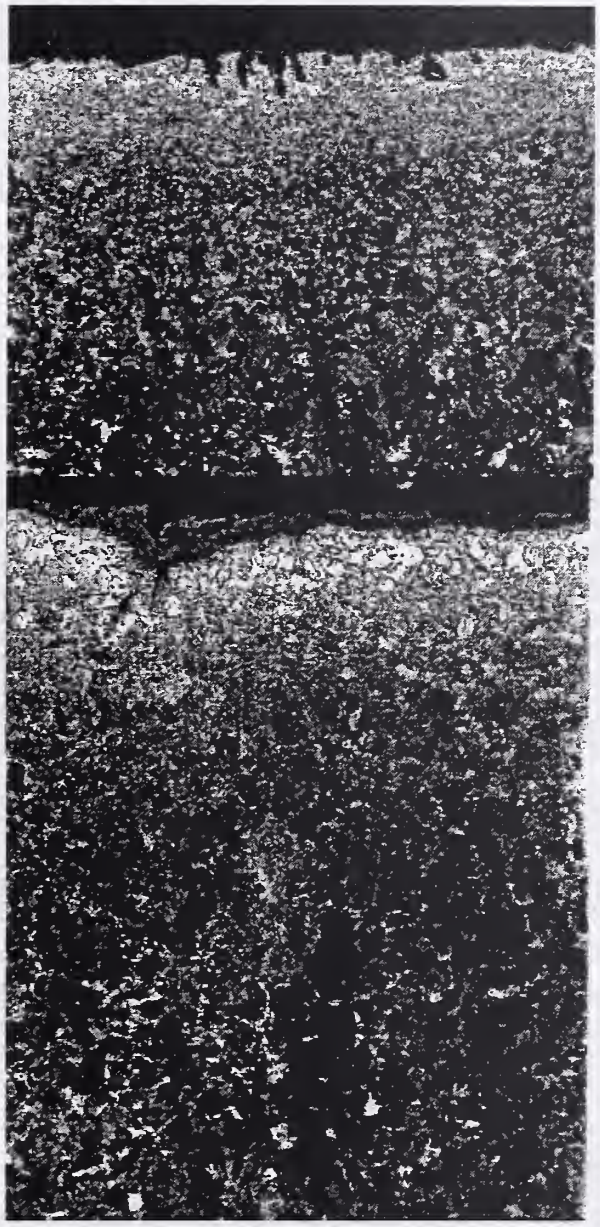


Figure 5. Photomicrographs of "long bar." Note the lesser amount of decarb layer compared to the "short bar," and the presence of the primary microstructure. Mag: (a) 50 X; (b) 100 X; Etch: 2 % Nital

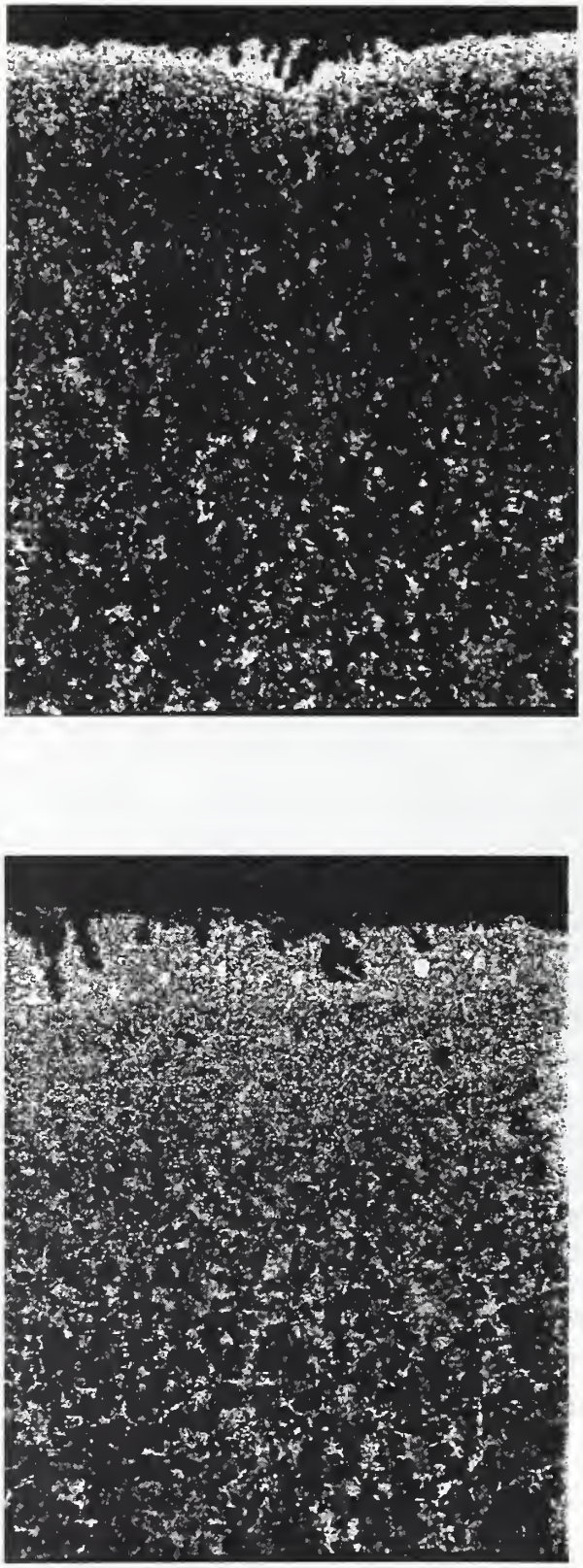


Figure 6. Photomicrographs of the "long bar" showing another region of the decarb and primary microstructure.
Mag: (a) 50 X; (b) 1OO X; Etch: 2 % Nital

Figure 7. Photomicrograph of the specimens with cracks around the specimen with the decarb layer. Note the absence of decarb layer in the cracked specimens.

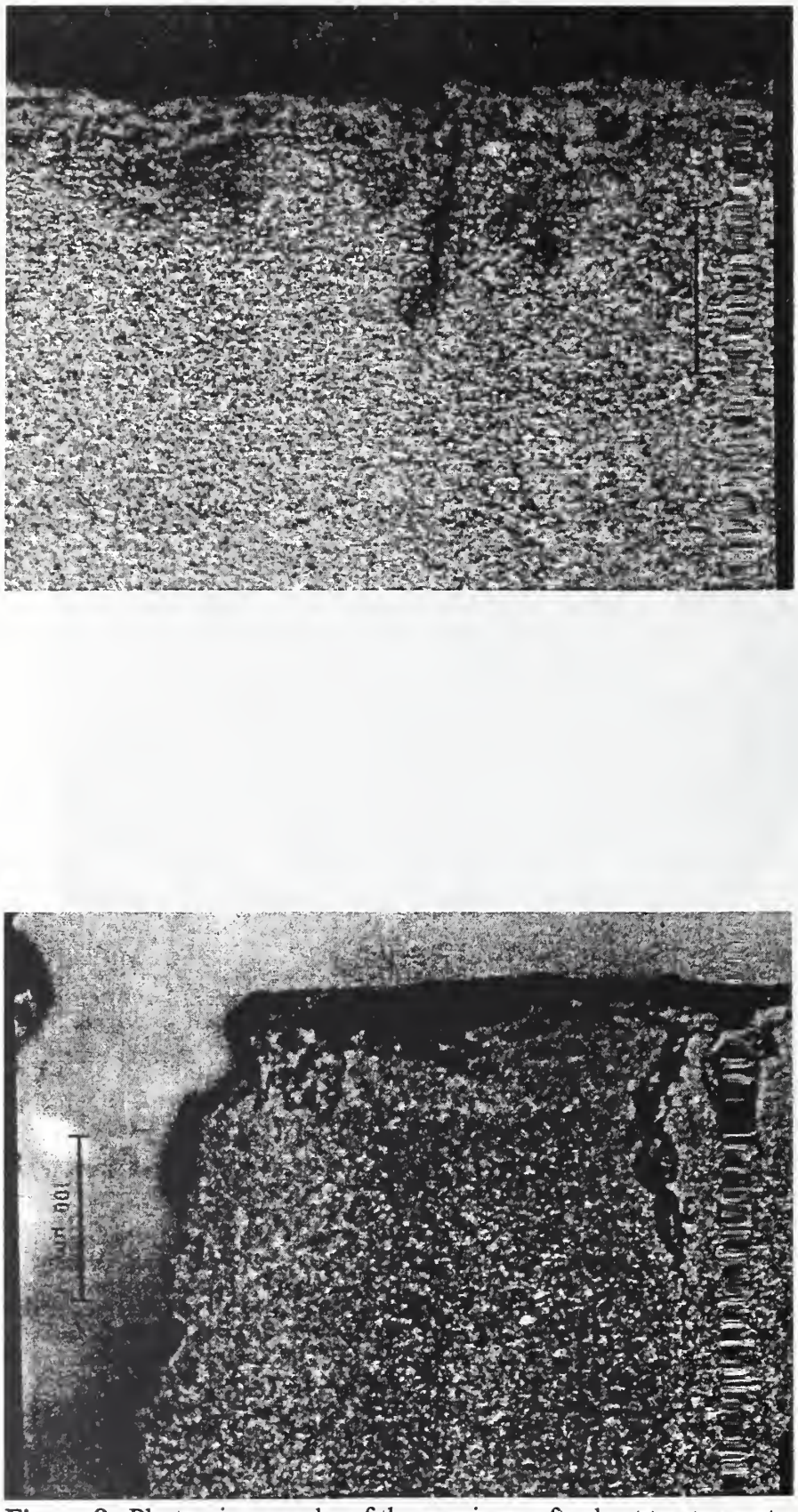
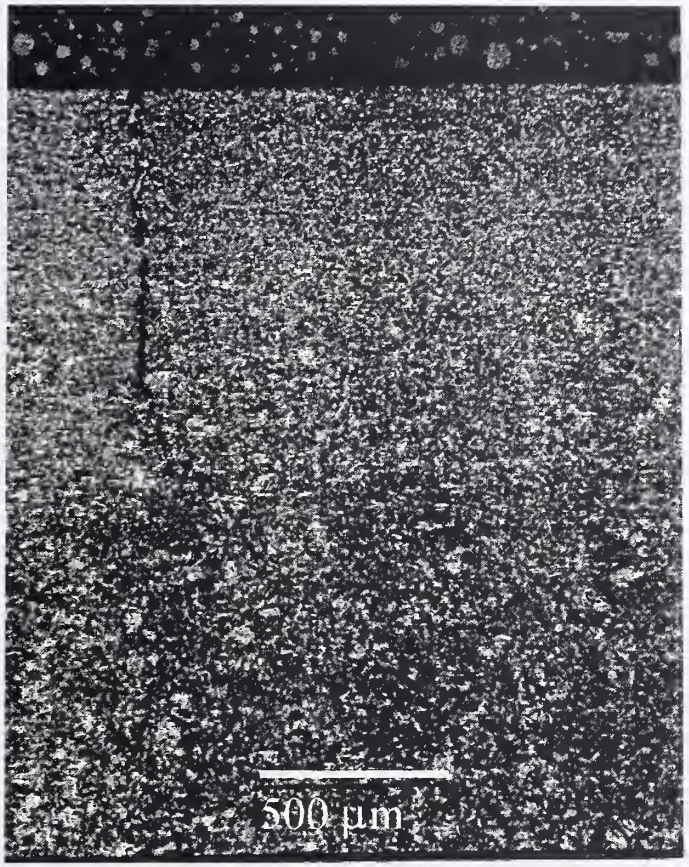


Figure 8. Photomicrographs of the specimen after heat treatment, showing the cracks.



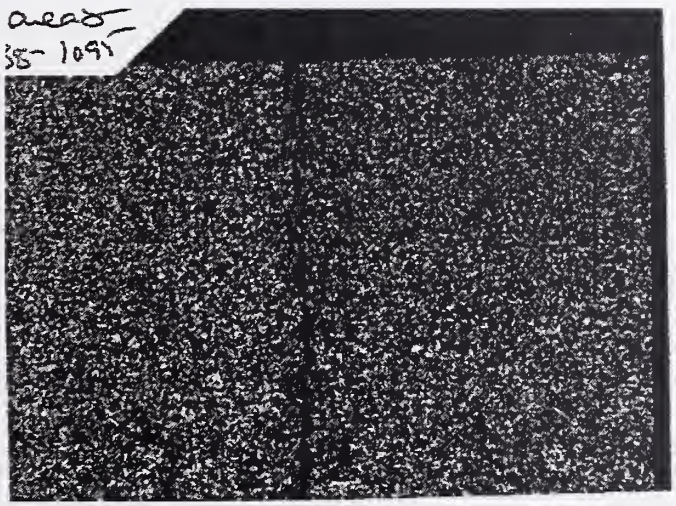


Figure 9. Photomicrographs of the specimens after heat treatment, showing the major crack in the specimens.

Part 3: Specimen, Machine, and Procedure Effects

A. K. Schmieder

Report of Measurement of Three Charpy Impact Machines at The National Institute of Standards and Technology

Introduction

During the period from September 21, 1992, through September 25, 1992, measurements of force and distance were made by instruments with calibrations traceable to national standards.

In general, the measuring procedures were those prescribed by ASTM - Test Method E23-91a. Section numbers in this report all refer to this test method.

This study has two primary objectives. The first primary objective was to measure the frictional losses by the change in elevation. The second primary objective was to evaluate the scale errors at the absorbed energy levels of the NIST verification specimens. Frictional losses were measured by the change in elevation during a single free swing, as described in Section 5.2.6, and also by reading the scale after 11 half swings, as described in 5.2.6.2. In order to obtain loss per swing, the accumulated loss is divided by 10 rather than by 11, as stated in 5.2.6.2. An additional test, not described in 5.2.6.2, was made during which the pointer is reset after the first, third, fifth, seventh, and ninth swings. The difference in scale readings for the two tests divided by five is an estimate of the loss due to the pointer friction. The results of these measurements of friction are shown in **Table 1**.

To evaluate the scale errors at the absorbed energy levels of the NIST verification specimens, the correction for friction loss was made by each of the two methods recognized by ASTM E23; namely, loss independent of angle of swing (Section 5.2.7), and loss proportional to the angle of swing (Section 9.1.1). These results are shown in **Table 2**.

The Satec machine and the Tinius Olsen machine have scales graduated in foot-pounds of energy absorbed by the specimen being broken. The indicated reading is simply the reading of the pointer position along this scale. The Tokio Koki machine has a protractor scale, so the observed value was the angle of rise of the pendulum after breaking the specimen. This angle was converted to energy in foot-pounds by a conversion table furnished by the manufacturer of the machine. The difference between the energy values calculated from the calibrating instruments and the difference calculated from the conversion table was several times as great as the corresponding values for the other two machines. To identify the source of these differences, the angles measured by the protractor readings were compared to those from the extensometer, and also the conversion calculations were repeated using various values of the physical dimensions of the machine. The results of these comparisons are shown in **Table 3**.

The individual values used to calculate the results in **Tables 1, 2, and 3** are listed in **Table 4**. A discussion of some of the notable results follows each table. The next section of this report shows an estimate of the accuracy of the measurements made by the calibrating equipment. The last two sections analyze in more detail observations listed earlier.

Table 1a. Frictional losses in percent of rated energy.

Machine Manufacturer	Satec	Tinius Olsen	Tokio Koki
Serial Number	1262	130005	878303
Rated Energy, ft.lb.	240	264	265
Frictional Loss in Pendulum and Pointer			
Elf Elevation Measurements, %	0.18	0.34	0.20
By Successive Swings, %	0.25	0.33	0.15
Friction Loss in Pointer Only			
By Repeated Swings Without Reset, %	0.00[1]	0.04	0.04
Elf Successive Swings With Reset , %	0.01	0.03	0.05

Table 1b. Comparison of frictional loss measurements between previous and present investigations.

	This Report		
	SATEC	TINIUS OLSEN	TOKIO KOKI
Pendulum Only, Single Swing , %	0.17	0.31	0.15
Pendulum Only, Successive Swings, %	0.24	0.30	0.10
Pointer Only, Successive Swings, %	0.01	0.03	0.05
	Earlier Report		
	Same Model	Same Model	Same Machine[2]
Pendulum only, Single Swing, %	0.30	0.25	0.10
Pendulum Only, Successive Swings, %	0.38	0.32	0.12
Pointer Only, Successive Swings, %	0.03	0.26	0.01

[1] It was observed that the pointer moved but fell back to its original position.

[2] Moved from Watertown, MA, to Boulder, CO, between tests.

Table 2. Percent error[3](a) in absorbed energy.

For SATEC Machine				
Indicated Absorbed Energy (ft-lbf)	10	12.5	75	150
Error, Assuming Friction Losses are:				
Independent of Angle of swing[4] (%)	0.01	-0.04	0.17	0.27
	(0.20)	(-0.72)	(0.53)	(0.43)
Proportional to Angle of Swing[5] (%)	0.01	-0.04	0.15	0.23
	(0.20)	(-0.80)	(0.47)	(0.37)
For TINIUS OLSEN Machine				
Indicate Absorbed Energy (ft-lb)	10	12	75	150
Error, Assuming Friction Losses are:				
Independent of Angle of Swing (%)	-0.00	-0.01	0.03	0.14
	(-0.10)	(-0.24)	(0.11)	(0.25)
Proportional to Angle of Swing (%)	-0.00	-0 .02	-0.00	0.07
	(-0.20)	(-0.40)	(-0.01)	(0.12)
For TOKIO KOKI Machine				
Indicated Absorbed Energy[6] (ft-lb)	10.01	12.71	71.85	150.1
Error, Assuming Friction Losses Are:				
Independent of Angle of Swing (%)	-0.11	-0.15	-0.38	-0.70
	(-2.80)	(-3.15)	(-1.41)	(-1.25)
Proportional to Angle of Swing (%)	-0.11	- 0.15	-0.40	-0.76
	(-2.90)	(-3.22)	(-0.56)	(-1.32)

[3] Error is taken as the indicated value minus the value calculated from the direct verification. Values without parenthesis are percent of rated energy, with parenthesis, of indicated energy.

[4] Friction loss assumed constant at all angles of swing.

[5] Friction loss assumed proportional to the sum of the angle of fall plus the angle of rise.

[6] Indicated absorbed energy is determined by reading the angle of rise on the protractor and then consulting the table furnished by the manufacturer of the machine.

Table 3. Percent error for Tokio Koki machine using various assumptions to determine indicated absorbed energy.

Values Shown in Conversion Table					
Scale Reading, Degrees Rise	110° 40'	107° 30'	106° 40'	89° 0'	65° 10'
Absorbed Energy by Specimen, (ft-lbf)	--	10.01	12.71	71.85	150.10
Absorbed Energy Determined by Calibration and Error					
From Calibration (ft-lb)	0.0	10.29	13.11	72.86	151.5
Error, Percent of Rated Energy (%)	--	0.11	-0.15	-0.33	-0.70
Error Using Indicated Energy from Recalculated Table					
Error, Percent of Rated Energy (%)	0.00	0.00	-0.04	-0.10	-0.19
Indicated Energy Measured Force and Length					
Error, Percent of Rated Energy (%)	0.00	0.01	-0.02	0.02	0.05
Error, Percent of Indicated Energy (%)	0.00	0.10	-0.05	0.05	0.08

Discussion of Friction Measurements in Table 1

Section 4.1.4 requires that for a single swing in the forward direction, the friction loss shall not exceed 0.75 % of the scale range and that due to the indicating devices, shall not exceed 0.25 %. The largest values in **Table 1a** were less than one half of these maxima.

Section 5.2.6.2 requires that for the successive-swing procedure, the loss per swing shall be less than 0.4 % of the scale range capacity. This corresponds to the values listed on the second line times 10/11. This requirement is also met by a generous margin.

A previous investigation[7] included machines of the same design as the SATEC and TINIUS OLSEN and the identical TOKIO KOKI machine. The frictional loss measurements of that investigation compare to this as follows:

The last two columns show closer agreement between the two reports for the successive-swing method than for the single-swing method. In the first column, the early report tested an older machine. If the 0.14 % difference between machines is assumed to be due to dirty bearings, then the results for the two Satec machines are consistent with each other and with the other two designs.

The conspicuous anomaly is the high ratio of the two values in the upper right hand corner of **Table 1**. This will be analyzed after the individual measurements are presented in **Table 4**.

---

[7] Schmieder, A.K. , Comparison of Metrological Techniques for Charpy Impact Machine Verification. Charpy Impact Test - Factors and Variables, ASTM STP 1072, 1990, pp. 25

The measurement of pointer friction without resets shows 0.04 % values. This is the smallest value that can be discerned on the scale. It corresponds to a pointer movement of about 0.01 inch. The successive-swing measurements of both pendulum and pointer friction have better precision because of the large quantity measured. The comparison above also indicates that they have better reproducibility. Therefore, they were used during the calculation of scale errors.

The pointer fall-back that was observed during operation of the Satec machine (see **Table 1**, footnote) indicates that it would be advisable to increase the friction drag on the pointer.

Discussion of Scale Errors Shown in Table 2

Table 2 shows the error as percent of the latched energy (rated capacity) and also as percent of the indicated reading. Section 4.1.3 requires that these errors shall not exceed 0.2 % of the former or 0.4 % of the latter, whichever is larger. With the exceptions noted, the errors at all levels met these requirements.

1. For the Satec machine, the error was slightly greater than the permitted amount at 150 ft-lbf when the friction loss was assumed to be independent of the angle of swing. When the friction loss was assumed proportional to the angle of swing, the error was within the required limits.

2. For the Tinius Olsen machine, all the errors were well within the acceptable amount for either of the two assumptions regarding distribution of function losses.

3. For the Tokio Koki machine, the error at the 72 and 150 ft-lbf levels was 2 to 3 times the permitted amount. The error was approximately proportional to the level when measured as a percentage of the range. This indicates a systematic error in the method of calculation. This possibility was analyzed quantitatively. The results are shown in **Table 3**.

Discussion of Table 3

The manual furnished with the machine gives the formulas used to calculate the conversion table together with an example of that calculation using hypothetical values for the measured quantities. The actual measurements for the machine calibrated are not stated. The manual also describes the method of treating frictional losses. The frictional loss of the pendulum was determined by a successive-swing method described in Section 5.2.6.2, except that the total loss was divided by 10 instead of 11. The divisor 10 is theoretically correct and is used in this report. (Note Section 5.2.6.2 does not give a method of determining the frictional loss of the pendulum, but a method of determining when the bearings are to be cleaned. The divisor 11 is probably appropriate for that purpose.)

The T-K manual apportions the friction loss according to the angle of rise, while the method of measurement gives the total for both rise and fall. This apportionment is theoretically incorrect. However, the effect on the calibration error is not nearly sufficient to account for the values observed.

The pointer friction is determined as described in Section 5.2.6.1 and is proportioned according to the angle of rise. This is theoretically correct. The method of Section 9.1.1, which apportions pointer friction according to the angle of swing, is theoretically incorrect. However, the small magnitude of the pointer's frictional loss makes this theoretical error insignificant.

The error shown in Section B of **Table 3** is copied from **Table 2**. The calibration error is determined as for the other machines. The absorbed energy is calculated from the direct measurement of force and distance. This value of absorbed energy is subtracted from the value in the conversion table for that angle of rise.

It is not known what friction loss was used to calculate the furnished table. For the recalculated table, the measured loss of 0.15 % of range was used. The example shown in the manual (not a case from the table furnished) used a friction loss of 0.35 % of range. If this is approximately the value used for the furnished table, that could cause a significant portion of the calibration error shown in Section B.

Section C of **Table 3** shows that when the conversion tables are recalculated by the method described in the machine manual but with the measured value of friction loss, the calibration errors are within the permitted limits. The difference between the furnished table and the recalculated table appear not to be primarily arithmetical. However, a small arithmetic error is obvious in that the furnished table value for absorbed energy at an angle equal to the angle of rise for a free swing is not zero, as required by the algebraic formulas. If it is not an arithmetic error, the difference in the two tables must be due to the input quantities, of pendulum length, weight, or friction loss.

The pendulum is inscribed with its weight and the distance from the axis of rotation to the center of gravity. Presumably, these values were used in calculating the furnished table. As shown in **Table 4**, the values measured during the direct verification were slightly different. Using the measured values results in the additional reductions in calibration error shown in Section D of **Table 3**. This is not evidence that the measured values are more accurate than those inscribed. It is simply a mathematical result of the same values being used for the calculated reference values and for calculation of the conversion table.

It is recommended that:

a. The measurement of friction values be confirmed by several different observers using the successive-swing method.

b. The weight and length of the pendulum be confirmed using different calibrated instruments of the highest available precision.

c. Using these confirmed values and both E28 methods, calculate two new conversion tables. List all formulas and input values used.

Table 4. Measured quantities and intermediate calculated results.

Machine			SATEC	TINIUS OLSEN	TOKIO KOKI	
Length of Pendulum[8], inches			31.493	35.450	35.393	
Supporting Force, lb.			53.35	59.96	66.64	
Measured Latched Angle, degrees			134.91	119.13	120.74	
Latched Energy, ft-lbf			238.86	263.37	266.09	
Friction Loss, ft-lbf			0.38	0.86	0.53	
Angles of Rise at:						
					Clinometer Reading	Protractor Reading
151	or[9]	150.10 ft.lb.	68.69°	68.73°	65° 2.5'	65° 10'
75	or	71.85	80.27	93.38	89° 53.5'	89° 0'
12.5	or	12.71	127.83	114.28	106° 32.5'	106° 40'
10	or	10.01	129.20	115.18	107° 23.5'	107° 30'
0	or	0.0	134.69	118.82	110° 34.5'	110° 40'
Scale Reading after 11 swings:						
Without pointer reset			5.9 ft.lb.	8.6 ft.lb		109° 48'
With 5 pointer resets			6.0 ft.lb.	9.0 ft.lb.		109° 38'

[8] Distance from center of rotation to center of strike.

[9] First value applies to Satec and Tinius Olsen, second to Tokio Koki machine.

Discussion of Table 4

The primary purpose of including **Table 4** is to facilitate confirmation of the accuracy of the calculated results and comparison of successive calibrations. In addition, the table reveals an interesting relationship between the angle as measured by the protractor attached to the T-K machine and as measured by the clinometer.

The last column shows that the two measurements of angle of rise differ by 5.5 to 7.5 min throughout the 65° range. Since the protractor is calibrated in 10 min intervals and is readable to ±1 min, the range of difference is not significant and the value of 6.5 min may be used as a constant bias.

The protractor read zero when the striker rested against a 10 mm diameter pin pressed against the anvils. The clinometer readings were adjusted to read zero in this pendulum position also. Sources of this bias include:

a. An inaccuracy in the zero-degree graduation.
b. An error in measuring the adjustment quantity for the clinometer.
c. A calculation error.

Estimates of Accuracy of Calibration Equipment

1. Balance Used to Determine Weight

To calibrate the balance, 0.5 lb. lead weights were hung in place of the pendulum. These were matched to a certified standard weight, using a precision scale reading in 0.1 gr intervals. The 10-lb poise and a graduated precision level on the beam allowed observation of the angle of swing of the beam during at least 10 cycles before coming to rest. This resulted in minimizing friction effects both in the balance and in the bearings of the pendulum.

The scale is graduated in intervals of 0.01 1b. Repeatability is also 0.01 lb. A known source of inaccuracy is the variability in the amount of adhesive putty used to attach a 10 mm diameter pin to the striker at the center of strike. Estimated inaccuracy is ±0.03 lb or ±0.05 %. Section 5.2.3.3 requires that the measurement of force be accurate within 0.4 %.

2. Pins and Vernier Caliper Used to Measure Lengths

A calibrated 18 in caliper certified to be accurate to 0.001 in was used to measure the length of rods bolted together to measure the pendulum length. The caliper is graduated in 0.001 in divisions. Seven interfaces including the ends occur in the assembled rod system so the accuracy is estimated to be ±0.007 in, or 0.02 %. The repeatability of these measurements was ±0.002 in.

3. Clinometer Used to Measure Angles

The clinometer is graduated in intervals of one minute of arc. The guaranteed accuracy is 1.5 min. In the range of angles used for these measurements, the instrument was verified by

measurement on a typical machine using a precision level to establish a reference plane and a calibrated ruler, graduated in 1/64 in intervals, to measure height. The two measurements of the angle of the pendulum agreed within the readability of the ruler, that is, 0.05 %. The accuracy of the angle measurements is estimated to be 0.04 %.

Both the inaccuracy of the length measurement and the angle measurement contribute to the inaccuracy of the elevation measurement. Using the variance method of combining, the estimated overall error of measurement for elevation is

$$\pm\sqrt{(0.02^2 + 0.04^2)} = \pm 0.05\%.$$

Section 5.2.3.4 requires elevation measurements to 0.1 %.

4. Estimated Accuracy of Calibration Values for Absorbed Energy

Similarly, combining the contribution of the force measurement and the angle measurement gives

$$\pm\sqrt{(0.05^2 + 0.05^2)} = \pm 0.07\%.$$

Analysis of the Difference Between Protractor and Clinometer Readings

A simple procedure to evaluate the accuracy of the protractor is to measure the chord from 120° to 60° and compare it to that from 60 to 0°. If a graduation error is the source of the bias, the latter distance would be shorter than the former by two-thirds of the distance between graduation marks. The bias could also result from the circular arc of the protractor being eccentric to the axis of rotation of the pointer. This is easily detected by scribing a fine line on the pointer at the arc at zero and swinging the pointer through full scale while observing any shift of the line relative to the arc.

Unlike the other two machines, the pendulum stem of the T-K machine is tapered so the adjustment factor for the clinometer was several degrees. A test of the accuracy of this factor is to compare the friction loss calculated from a single swing to that calculated from successive swings. The former depends on only the clinometer readings and the latter is completely independent of the clinometer readings. The reference cited earlier shows the average ratio of

$$\pm\sqrt{0.02^2 + 0.04^2} = \pm 0.05\%$$

these two is 0.83. For the T-K machine, the ratio was 0.83. The ratio of the results in **Table 1** is 1.33. If an error in the zero reference of the clinometer gives a smaller angle of rise on the free swing (as shown in **Table 4**), a corresponding increase in the latched angle occurs. The apparent frictional loss and the ratio would both be increased from their correct value, as shown by the results. However, if the true value of the ratio was 0.83 and the loss was 0.15 %, the latched

$$\pm\sqrt{0.05^2 + 0.05^2} = \pm 0.07\%$$

angle would be 110.79° for an angle of rise of 110.67°. If a 6.5 min bias is assumed, these angles would become 110.90° and 110.56°, respectively, giving an apparent ratio of 2.73 as compared to the observed value of 1.33. Therefore, it appears that the bias of 6.5 min is not due to an error in the clinometer correction. If the observed bias is due to an error in the protractor scale near zero, the value of ratio would not be changed and the unexpectedly high value is left unexplained

Analysis of Unexpectedly High Ratio of Two Friction Measurements for the T-K Machine

The values in **Table 1** show that a ratio of the friction measured by a single swing to that measured by successive swings is 1.33 for the T-K machine. The average of the other machines in **Table 1** is 0.87. The average of eight machines tested earlier is 0.83. The T-K machine tested earlier also had a value of 0.83. In the interval between the two tests, the machine was moved. The striker is now magnetized. This occurs when some hardened steels are repeatedly struck by a blow causing bending. If the machine was originally installed new and not magnetized, the pendulum would hang vertically and the protractor could be set to a true zero. During the later installation, the magnetized striker would tend to move the pendulum toward the anvils giving a false zero. Calculations show that if the true ratio were 0.83, a horizontal force of approximately 2 oz toward the anvils would be sufficient to result in an observed ratio of 1.33.

Most steel specimens are held in a non-magnetic chuck while grinding. Residual magnetism might also cause a deviation while the pendulum is being leveled so the scale reads 0 degrees when the striker is almost touching a steel specimen. During this calibration, a gap of 0.010 in was observed.

Daniel P. Vigliotti,[1] Tom A. Siewert,[1] and Chris N. McCowan[1]

Maintaining the Accuracy of Charpy Impact Machine*

Reference: Vigliotti, D. P., Siewert, T. A., and McCowan, C. N., "**Maintaining the Accuracy of Charpy Impact Machines,**" *Pendulum Impact Testing: A Century of Progress, ASTM STP 1380*, T. A. Siewert and M. P. Manahan, Sr., Eds., American Society for Testing and Materials, West Conshohocken, PA, 1999.

Abstract: The quality of the data developed by impact machines tends to degrade over time, due to the effects of wear and vibration that are inherent in the test. This is the reason that impact standards specify periodic direct and indirect verification tests. Each year, we provide reference specimens for indirect verification of over 800 machines around the world. From evaluation of the absorbed energies and the fractured specimens, we are able to deduce the origin of energies that are outside the allowed ranges, and report these observations back to the machine owners. This report summarizes the basis for these observations and will allow, it is hoped, machines to be maintained at higher levels of accuracy.

Keywords: absorbed energy, Charpy V-notch, impact test, machine repair, misalignment, verification testing, worn anvils

Introduction

The low cost and simple configuration of the Charpy impact test have made it a common requirement in codes for critical structures such as pressure vessels and bridges. However, accurate results can be obtained only from impact machines that remain in good working condition, such as within the tolerances specified by ASTM Standard Test Methods for Notched Bar Impact Testing of Metallic Materials (E 23). We find that many of the critical tolerances can be monitored by post-fracture examination of the standardized verification specimens that we distribute.

Our examination of over 2000 sets of these specimens each year allows us to identify problems that are often not recognized during routine measurement of machine dimensions or routine check procedures. We have learned to recognize what marks on the broken verification specimens indicate factors that could be affecting the results. We can then advise our customers to recheck or replace the anvils or the striker, tighten bolts, check bearings, check machine alignment or level, check cooling bath or thermometer, or review testing procedures. This paper describes the most common problems that we detect, and gives advice on how to avoid or correct most of them.

---

[1] Technician, Group Leader, and Metallurgist, respectively, Materials Reliability Division, National Institute of Standards and Technology, Boulder, Colorado 80303

* Contribution of NIST; not subject to copyright.

Direct Evaluation

A routine check consists of a free swing check and a friction and windage check. The free swing is a quick and simple test to determine if the dial or readout is performing accurately. A proper zero reading after one swing from the latched position is required on a machine that is equipped with a compensated dial. Some machines are equipped with a non-compensated dial. Such a dial is one on which the indicator cannot be adjusted to read zero after one free swing. The user should understand the procedure for dealing with a non-compensated dial. This information should be available from the manufacturer.

The friction and windage test will give the user the condition of the bearings. We suggest that any deviation of more than 5% is excessive and the bearings should be inspected.

We suggest that the user develop a daily log or shift log to be kept with the machine. The log can be used to track the zero and friction values. The log can also include information such as number of tests, materials tested, and any other useful comments. A sample log is attached as Appendix 1.

Machine Preparation

The Charpy test is a dynamic test. Therefore, bolts may loosen over time. The tightness should be checked on the anvil bolts, the striker bolts, and the baseplate bolts. The manufacturer can supply the torque values for the anvil and striker bolts. The base-plate bolts should be torqued to the recommended torque values for the grade and size of the nuts and bolts. We recommend the use of "J" or "T" bolts only. (See Appendix 2.) We do not recommend lag type bolts. These bolts are made to withstand only static loads. We believe that over time, the insert portion of lag bolts will loosen in the concrete. As lag bolts are continually tightened, they can pull out of the concrete and be tightened against the base of the machine, giving the impression of a properly mounted machine. This condition is very difficult to detect. A machine with this problem will cause high-energy values at the low-energy level. The procedure used to mount the master reference machines is attached as Appendix 2.

The anvil and striker radii should be carefully inspected for proper dimensions and for damage. Damage can be detected easily with a visual inspection and a check for smoothness by running a finger over the radii. We find that radius gages are usually inadequate to measure the critical radii. We recommend making molds of the radii and measuring the molds on an optical comparator. Occasionally even a new set of anvils and striker may have incorrect radii. We recommend that new anvils and strikers be inspected before being installed in the machine. Since the radii can be considered consistent before use, they can be measured directly on an optical comparator or other optical measurement system.

We recommend centering tongs such as those described in ASTM Standard E 23. The tongs should be inspected for wear or damage. A proper set of tongs is critical for the accurate placement of the specimen. Some machines are equipped with a centering device. The device should be inspected for wear and proper operation. We do not recommend the use of centering devices for low temperature testing because it can delay the time between removal from the bath and fracture, and so may exceed the allowed five-second interval.

The temperature indicator should be calibrated immediately before testing. Since ice water and dry ice have constant temperatures, they make quick and easy calibration media.

Post-Fracture Examination

Many machine problems can be monitored by post-fracture examination of the NIST standardized verification specimens. Following are the most common of these problems. In many cases, suggestions on how to correct or avoid them in the future are included.

Worn Anvils - Most of the wear of an impact test machine occurs on the anvils and striker. We evaluate this wear by examining the gouge marks that are formed on the sides of high-energy specimens when they are forced through the anvils. Anvils that are within the required tolerance of the standard will make a thin, even gouge mark all the way across both pieces of the broken specimen. As the anvils wear, they will make a wider, smeared mark across the specimen halves. **Figure 1** shows the change in the gouge marks. When the wider smeared marks are observed on a customer's specimens, we recommend that the anvils be changed, because the reduction in energy needed to push the specimens through worn anvils eventually drops the machine below the lower tolerance in the energy range. You can monitor the wear on your machine by retaining some specimens that are tested with new anvils and comparing them to specimens of similar composition and hardness that are tested as the anvils wear. For specimens at a similar absorbed energy, the gouge marks will grow wider and smoother as the anvils wear.

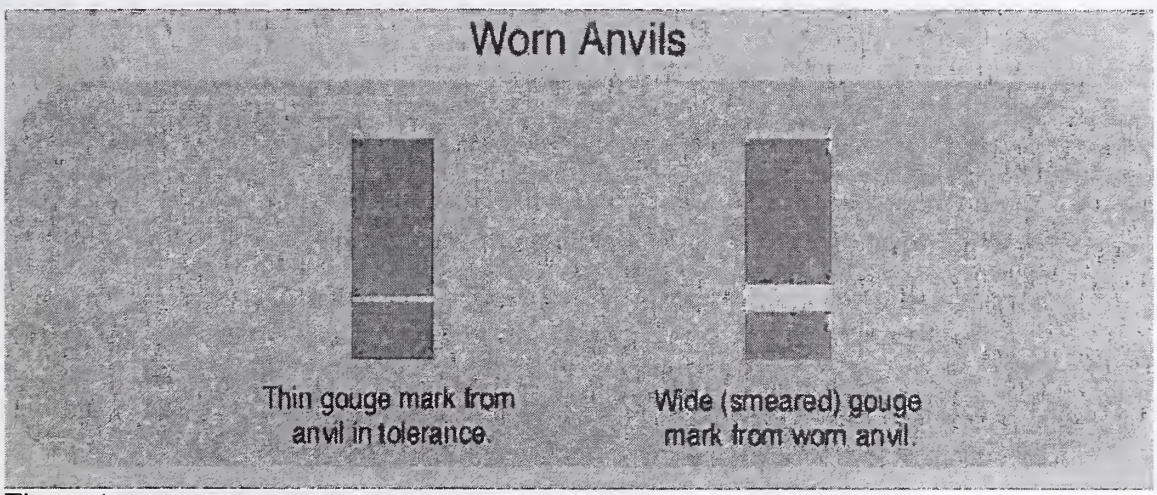


Figure 1

Off-Center Specimen - An off-center specimen strike occurs when a specimen is not centered against the anvils, so the striker contacts the specimen to the side of the notch. The low-energy specimen best indicates when an off-center strike occurs. We identify this condition on the specimens by finding that the gouge marks caused by the anvils are not equidistant from the machined notch edges, and the striker gouge mark is offset the same amount from the notch (**Figure 2**). Also, as seen in **Figure 2**, the fracture surface of a correctly tested low energy specimen is flat and both halves are even. However, the fracture surfaces of a specimen that has been tested off-center are on an

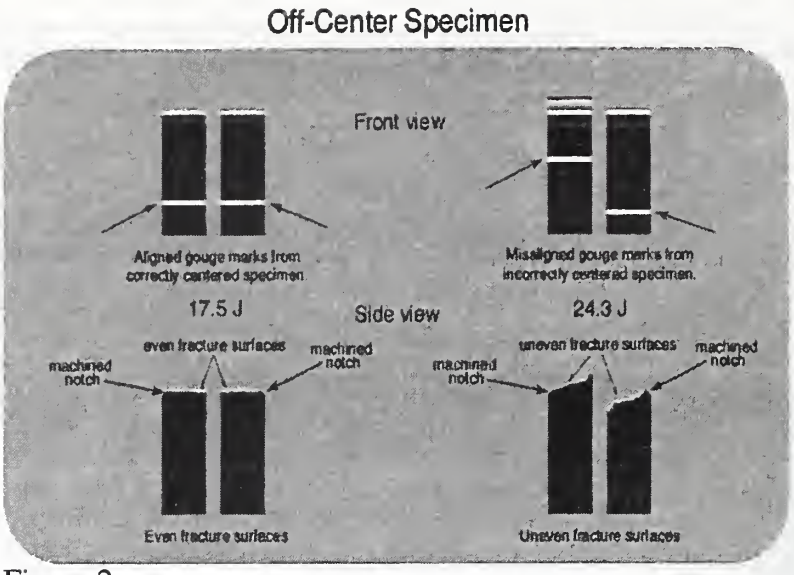


Figure 2

angle. The more off-center the strike, the steeper the angle will be. This problem increases the energy needed to fracture a specimen. The most common causes for this slipping are worn or damaged centering tongs, a worn or misaligned machine centering

device, careless test procedures, or the use of a cooling fluid that is too viscous at the test temperature, which causes the specimen to float on the specimen supports. Most machine manufacturers should be able to provide new centering tongs. We have found that ethyl alcohol is one of the best cooling media because it seems to evaporate quickly from the bottom of the specimen to prevent specimen floating.

Off -Center Striker - This differs from the off-center specimen in that the specimen is centered against the anvils so the anvil gouge marks are equidistant from machined notch edges. However, the striker does not contact the specimen precisely opposite the notch. **Figure 3** shows this appearance. An off-center striker is usually

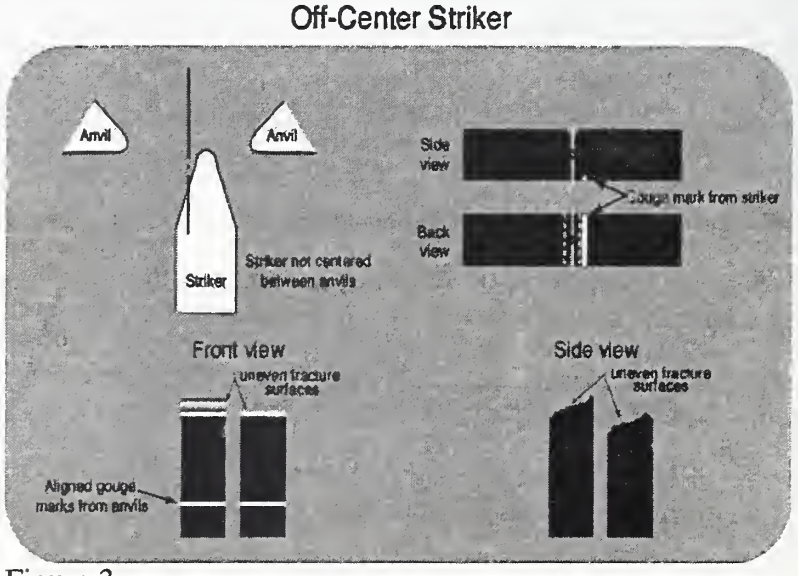


Figure 3

attributed to the pendulum shaft shifting off center. This shift can be the result of a loose alignment ring on the shaft or a loose bearing block on the machine. This problem also increases the energy needed to fracture specimens at all energy levels.

Uneven Anvil Marks - Frequent testing of subsize specimens can cause the anvils to wear unevenly. **Figure 4** shows an example of these uneven wear marks at each energy level of our reference specimens. Since this wear is restricted to a small area that the full-size reference specimen contacts, there is usually no effect on the energy required to fracture the specimen. This anvil condition presents two problems. First, since subsize wear is usually not indicated by a change in the energy required to break a reference specimen, inspection of the broken specimen is required. This wear will cause the anvils to be out of tolerance according to the requirements in the standard. This means that the machine does not meet the direct verification requirements of the standard and is therefore not eligible for the indirect verification process. The second, and more important problem, is that the subsize specimens are being tested in an area of the anvil that is worn. When the wear is substantial, this condition will produce artificially low sub-size energy values. The anvils should be replaced on a machine with this condition.

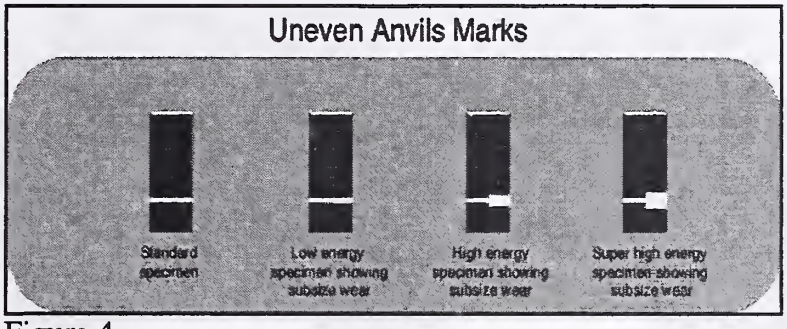


Figure 4

Chipped Anvils - Sometimes an anvil can be chipped. **Figure 5** shows that this condition can be detected easily on all three energy reference specimens. The low-energy specimen is affected the least amount because it is the hardest specimen and therefore has a more brittle fracture. The ductile high-energy specimen will produce higher than normal energy results and the very ductile super-high-energy specimens are affected most by a chipped anvil. This condition should be detected easily by a visual inspection before using the machine. New anvils are required when an anvil is chipped.

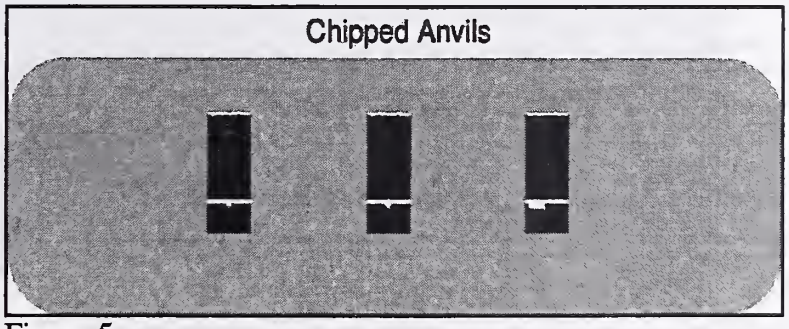


Figure 5

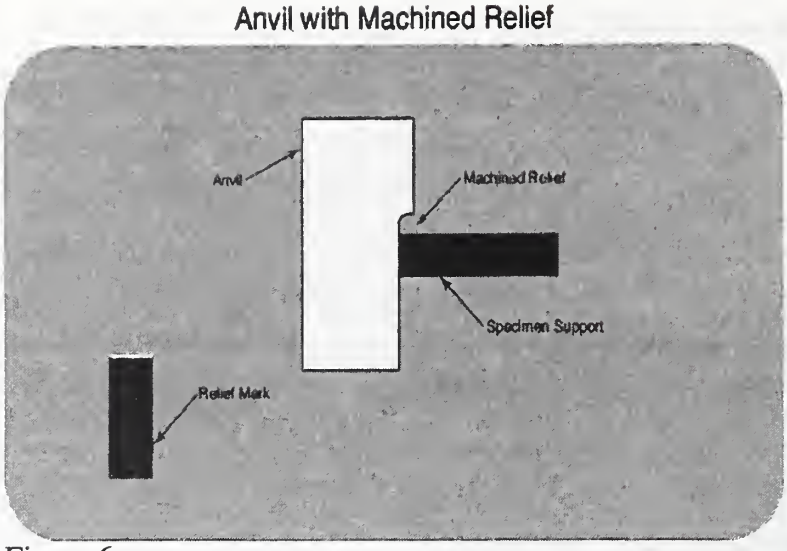


Figure 6

Anvil Relief - Some Charpy machine manufacturers have designed a machined relief at the bottom of the anvil (**Figure 6**). This anvil design does not meet the direct verification requirements of ASTM Standard E 23. The relief has caused high-energy results in our ductile high and super-high-energy specimens. It can also cause twisting of the specimens during fracture that may also contribute to energy values higher than normal at all energy levels. Since the relief is designed into the anvils and does not appear to add an excessive amount of energy to the test, we at NIST continue to verify these machines.

Damaged Anvils - Under some test conditions, usually for elevated temperature testing, the anvils can wear to a rough finish that creates excessive friction (**Figure 7**). This damaged condition is detected best on the high and super-high specimens. Damaged anvils usually cause the gouge marks to become wider and push the specimen material to form a ridge that can easily be detected with the fingernail. This damage usually causes artificially high-energy results at the high and super-high energy levels. Damaged anvils must be replaced.

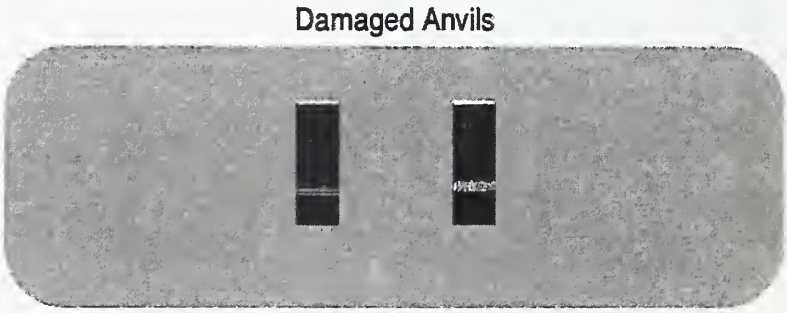


Figure 7

Bent Pendulum - **Figure 8** shows the gouge marks created by a pendulum bent in the direction of the swing. This gouge mark is usually deeper on the top edge of the specimen as it sits in the machine. As shown in **Figure 9**, the striker contacts the top edge of the specimen first, causing excessive tumbling and twisting. This excessive activity can cause the specimen to interact with the striker or the pendulum after fracture to create additional energy loss. A bent pendulum can be detected by placing an unbroken reference specimen in the machine and placing a piece of carbon paper on the surface opposite the notch. At this point, lightly tap the striker against the specimen. This will make a mark on the specimen that can be inspected. If the pendulum is not bent, the mark should appear the same width across the specimen. If the pendulum is bent, the mark will be wider at one edge and become thinner or even not visible at the other edge (**Figure 8**). We recommend that a new pendulum be installed on a machine with this problem.

Figure 8

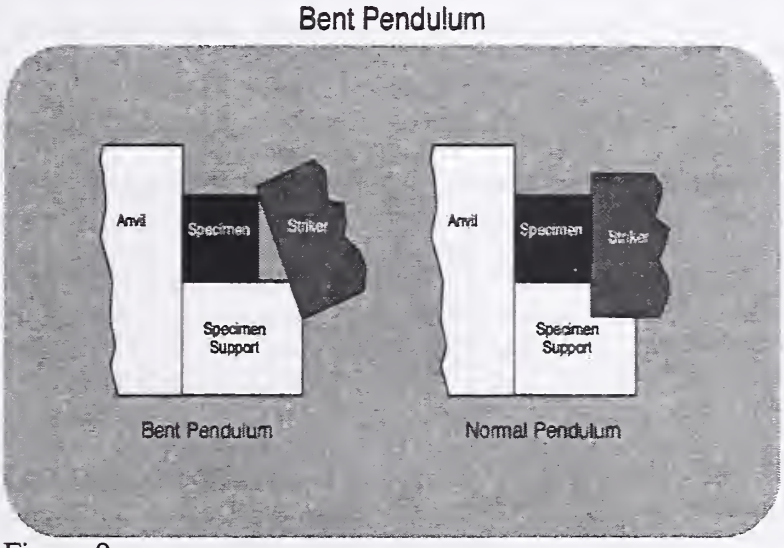


Figure 9

Summary

The condition and accuracy of Charpy machines cannot be checked only by comparing results of NIST reference specimens to the Master Reference Machines located at NIST, Boulder, CO. Some machine problems cause artificially low results while other machine problems cause artificially high results. In addition, deviations in procedures can cause similar results. These machine problems and procedural deviations may go undetected for years without some sort of physical check. For this reason, examination of the broken specimens is a critical part of the verification process. Many machine problems can be avoided or corrected with the information presented in this paper. Also, suggested changes in procedure can help can help to insure a successful test. To obtain verification specimens or to clarify procedures for verification testing, you may use the following information:

> Questions on verification procedures can be answered by the Charpy Program Coordinator. Phone: (303) 497-3351, fax: (303) 497-5939, or email: daniel.vigliotti@nist.gov

APPENDIX 1

EXAMPLE LOG

APPENDIX 2

MOUNTING PROCEDURE FOR REFERENCE CVN MACHINES

This is a detailed procedure developed by NIST to mount the three Master Charpy Reference Machines. This procedure is not intended to be substituted for any installation procedure provided by the manufacturer of the machine.

The foundation of the impact machine is critical to insure accurate results. Energy losses through the foundation must be kept to a minimum. We recommend making a foundation of 7000 pound mix concrete that measures 152.4 cm (60 in.) long by 91.4 cm (36 in.) wide by 45.7 cm (18in.) thick. Usually you will need to cut a hole in the floor to accommodate the new foundation. If other equipment in the area could affect the machine operation, you should isolate it from the floor with expansion-joint material.

Hold-down bolts used to secure the machine to the foundation should be of the inverted "T" or "J" type. The bolts, nuts, and washers should have a strength of grade 8 or higher. We recommend using bolts with a diameter of 22 mm (7/8 in.). At NIST we used 22 mm (7/8 in.) grade 8 threaded rod, cut into pieces 61 cm (24 in.) long. We then welded 22 mm (7/8 in.) pieces of the same threaded rod, six inches long, to the end of the 61 cm (24 in.) pieces to make inverted T bolts.

We then positioned the machine over the center of the foundation hole. The machine was held approximately 10.2 cm (4 in.) above the floor using spacers suitable to hold the weight of the machine. The T bolts were positioned in the machine-base mounting holes with a nut below and above the base of the machine. The nuts were tightened to keep the T bolts straight while the concrete was poured. The ends of the T bolts were positioned approximately 2.5 cm (1 in.) from the bottom of the hole. The machine was then leveled on the spacers. Leveling did not need to be as accurate as the final leveling. Reinforcement bars were attached to the top of the horizontal rod previously welded to the bottom of the T bolts. The reinforcement bars were attached in the form of a box connecting the four bolts. Another box formation of reinforcement rods was attached to the T bolts 25 cm (10 in.) above these rods. The concrete was then poured under the machine. The concrete was finished as level as possible at this time. Before the concrete fully hardened, we removed concrete from around each T bolt to create a cavity of approximately 2.5 cm (1 in.). This cavity would enable a nut to be threaded below the surface of the concrete. The machine was left in this position for 72 hours.

After 72 hours, the nuts on top of the base plate were removed and the machine was lifted off the T bolts. The bottom nuts were then threaded down into the cavities created before the concrete hardened. The nuts were left high enough on the T bolts to enable the use of an open-end wrench to adjust them after the machine was positioned on them. At this point, the base of the machine was coated with a light oil to keep grout from adhering to it. The machine was then lifted onto the T bolts and was positioned on the adjustment nuts. The machine was now ready to level. A machinist's level was used to insure meeting the tolerance of 3:1000 in. The critical leveling procedure was done using the four nuts under the machine. After the machine was leveled, we wrapped the outside of the nuts with duct-seal putty to facilitate their removal from the T bars later in the process.

At this point, the base of the machine was ready to grout. Heavy cardboard forms were placed around the base of the machine to keep the grout under the machine. It was necessary for the grout to flow completely under the machine, making sure the base of the machine was in total contact with the grout. The grout was installed under the machine.

The machine was left in this position for 72 hours.

After 72 hours, the machine was lifted off the T bolts. The grout was inspected for cavities and for surface contact with the bottom of the machine. The putty was removed form around the nuts. The nuts were removed from the T bolts. After removing all debris from the grout, the machine was lifted over the T bolts and rested on the grout. Washers and nuts were installed and tightened. The level was checked at this point. The T bolts were cut off to approximately 12.7 mm (1/2 in.) above the nuts. The nuts were torqued to 380 ft-lb. The final level was checked at this point.

NOTE:
Special non-shrinking grout is recommended. This grout is available at most industrial hardware stores.

If you have any questions concerning this procedure, please contact Daniel Vigliotti by phone at (303) 497-3351, by fax at (303) 497-5939, by email at daniel.vigliotti@nist.gov, or by mail at NIST, Division 853, 325 Broadway, Boulder, CO 80303-3328.

C.N. McCowan and D.P. Vigilotti

Effect of Reduced Notch Radius on Charpy Impact Energy

Introduction

Impact specimens produced for verification testing need to have the lowest possible variation in absorbed energy. This reduces the overall variation in the system, and allows for more certain estimates of the true mean energy for a given group of verification specimens. The system variation can be defined as:

$$\sigma^2 = \sigma_s^2 + \sigma_e^2 + \sigma_p^2, \quad (1)$$

where σ_s^2 is the variation due to inhomogeneities in the test specimens, σ_e^2 is the variation due to errors caused by measurement uncertainty and test equipment in general, and σ_p^2 represents the variations due to test procedures (operators, data collection, etc.). Steps have been taken to reduce many of the factors contributing to the system variation, by such means as: (1) a single operator used to perform all impact tests, (2) optical encoders implemented on each impact machine to remove operator bias and error from the data-collection process, (3) dimensional measurements of each verification specimen prior to testing, and (4) careful monitoring of room temperature and specimen temperature. Further improvements will most likely involve reducing specimen variation.

Considering the specimens, two types of inhomogeneities can readily be identified, dimensional and microstructural. As mentioned above, the dimensional variations of the specimens are controlled as closely as possible and the tolerances allowed for the specimens are as near as possible to a practical limit applicable to standard machining practices. The microstructural variation of specimens is minimized by controlling the chemistry and processing of the steel, and then by optimizing the heat-treating process used to produce specimens. But some variation due to microstructural inhomogeneities must be expected, and accepted. This is not to imply that there is no room for improved heat treatment procedures and alternate materials that might reduce variation. For a given material and heat treatment, however, one possible way to reduce the variation in absorbed energy is to change the specimen geometry in a way that helps to minimize the effects of microstructural variation. This report considers the effect of reducing the notch radius on the variation in impact energy.

Materials and Procedure

The impact specimens used for this study are verification specimens of 4340 steel that were heat-treated to produce a 100 % tempered martensitic structure with a hardness of approximately 45 HRC and an impact energy of approximately 16 J at room temperature. A control set of 25 specimens was machined to standard ASTM E23 Charpy impact specimen dimensions (0.25 mm notch radius, 0.01 in) and tested at −40 °C. A second group of 25 specimens was also machined to the standard 10 mm × 10 mm × 55 mm dimensions for an ASTM Charpy specimen, but the notch in these specimens was cut using an electrical-discharge machining (EDM) process with a wire, 0.25 mm (0.01 in) in diameter, resulting in a notch radius of 0.125 mm (0.005 in). The standard notch and the EDM notch respectively are shown in **Figures 1 and 2**. The straight notch of the EDM sample was grooved near the top to facilitate the use of centering tongs for impact testing.

Figure 1. Conventional/Standard notch for Charpy V-notched specimen.

Figure 2. Modified EDM notch .

Results and Discussion

The data for impact energy of the two groups of specimens are shown in **Figure 3** and listed in **Table 1**. The smaller notch radius significantly reduced the mean energy of the specimens. The mean energies for the modified EDM and standard specimen groups were 10.5 and 15.7 J, respectively. A portion of this decrease in energy might be argued to be a result of changes to the thin layer of steel at the tip of the modified EDM notch due to melting and resolidification (heat-affected zone). We did not attempt to separate this effect from that of the notch radius, because our purpose was to reduce scatter in the impact energy for the specimen, and not necessarily to determine the impact energy of the steel. Actually, we had hoped for a heat-affected zone at the EDM notch, which might aid in the crack initiation and reduce the scatter associated with the initiation event during fracture. A comparison between the coefficient of variation (CV) for the two groups of specimens, however, indicates that the notch did not reduce the scatter in the absorbed energy for the specimens. The CV of the EDM group, 0.099, is actually higher than the CV of the standard V-notch group, 0.068.

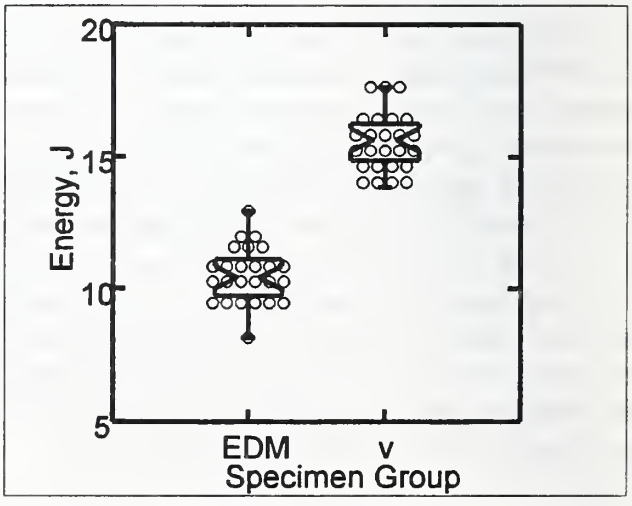


Figure 3. The absorbed energy data for the standard V-notch and modified EDM notched specimens. The data are boxed to contain 50 % of the values and notched at the median energy value.

There was some variation in the notch depths of the modified EDM samples measured prior to testing, as shown in **Figure 4**. A number of the specimens had notch depths well beyond the 0.025 mm tolerance allowed by ASTM Standard E 23. The variation in notch depth,

however, was not found to significantly affect the absorbed energy measured for the specimens. Removing the data for specimens outside the notch-depth tolerance (outliers) would decrease the scatter in absorbed energy for the modified EDM specimens, but not significantly.

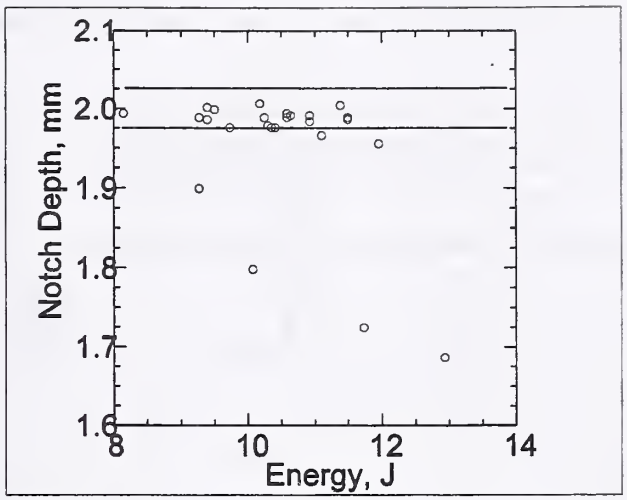


Figure 4. The notch depth versus the absorbed energy of the modified EDM specimens. Data within the band meet the ±0.025 mm tolerance for notch depth allowed for standard size Charpy impact specimens by ASTM E 23.

Follow-up Testing

To evaluate the effect of the EDM process on the absorbed energy of the low-energy specimens, additional tests were conducted. In these tests three types of notches were compared: (1) standard ground notch (45° angle with 0.25 mm notch radius, 0.01 in), (2) standard EDM notch (45 ° angle with 0.25 mm radius, 0.01in), (3) modified EDM notch (slot with 0.25 mm radius, 0.01 in). The notch radii were keep constant so that differences between the notch cutting process and notch geometry could be evaluated. In addition to low-energy samples, two groups of high-energy specimens were also tested.

The results of the testing are given in **Figures 5 and 6**, and **Table 2**. At low energy, no significant difference between the two notch geometries of EDM specimens was found. The EDM specimens with standard notch geometry had an average absorbed energy of 17.3 ± 0.7 J (12.8 ± 0.5 ft-lbf). The specimens with the modified notch geometry had an average absorbed energy of 16.8 ± 1.0 J (12.4 ± 0.7 ft-lbf). So, as might be expected, changing the notch from 45 ° to a slot does not influence the energy level of the specimens, while changing notch depth and radius does. The low-energy EDM specimens, however, were found to have higher energy than the low-energy specimens that had ground notches. The specimens with ground notches had an average energy of 14.8 ± 0.6 J (10.9 ± 0.4 ft-lbf). This would indicate that the EDM process changes the microstructure of the low-energy specimens (very near the notch) and this increases the absorbed energy of the specimens. At the high-energy level, this difference between EDM and ground notches was not apparent. The two groups of high-energy specimens tested had similar absorbed energies. The modified EDM notch specimens had an average energy of 93 ± 2.3 J (68.6 ± 1.7 ft-lbf). The high-energy specimens with ground notches had an average energy of 92.6 ± 2.4 J (68.3 ± 1.8 ft-lbf).

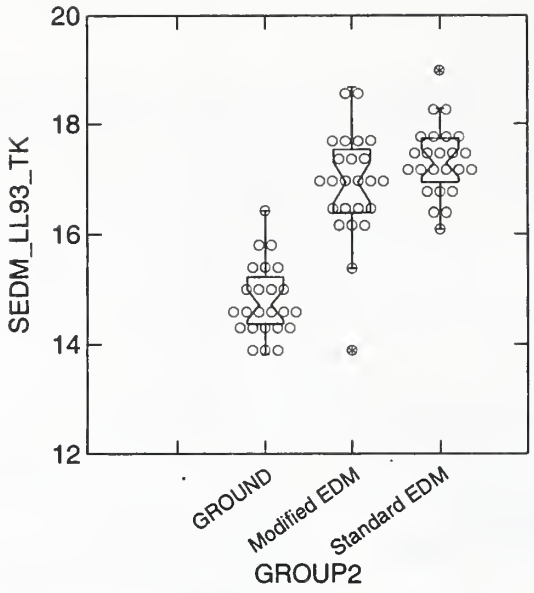


Figure 5. A group of standard ground specimens, a group of modified EDM specimens, and a group of standard EDM notched specimens. The specimens are low-energy specimens, LL93.

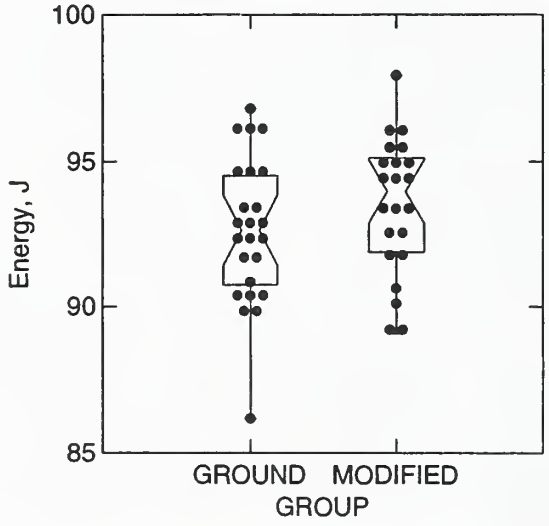


Figure 6. A group of modified EDM notched specimens and a group of standard, ground notched specimens. The specimens are high-energy specimens, HH86.

Table 1. Absorbed energies for specimen groups with standard ground V-notches and modified EDM notches.

Standard Ground V-notch Group		Modified EDM Notched Group	
Specimen ID	Absorbed Energy J (Ft-lbf)	Specimen ID	Absorbed Energy J (Ft-lbf)
396	14.79 (10.91)	491	12.93 (9.54)
40	15.65 (11.54)	548	9.27 (6.84)
1237	17.46(12.88)	378	9.74 (7.18)
27	15.80 (11.80)	1101	10.30 (7.60)
417	14.10 (10.40)	602	11.10 (8.19)
434	14.45 (10.66)	514	10.36 (7.64)
583	15.05 (11.10)	861	9.27 (6.84)
263	16.77 (12.37)	851	10.41 (7.68)
718	15.48 (11.42)	435	11.96 (8.82)
439	14.87 (10.97)	1004	8.15 (6.01)
744	14.20 (10.47)	1284	11.73 (8.65)
676	13.84 (10.21)	986	9.40 (6.93)
613	16.34 (12.05)	785	11.50 (8.48)
1013	16.00 (11.80)	929	10.07 (7.43)
108	15.31 (11.29)	1250	11.50 (8.48)
988	15.82 (11.67)	452	10.25 (7.56)
1087	16.60 (12.24)	856	10.64 (7.85)
381	14.87 (10.97)	358	10.93 (8.06)
273	15.31 (11.29)	66	9.40 (6.93)
876	14.87 (10.97)	441	10.59 (7.81)
158	16.08 (11.86)	813	10.59 (7.81)
1213	17.64 (13.01)	930	10.18 (7.51)
321	16.08 (11.86)	89	10.93 (8.06)
117	17.64 (13.01)	593	9.50 (7.01)
0	16.26 (11.99)	862	11.39 (8.40)
Mean Absorbed Energy =	15.66 J (11.55 ft-lbf)	Mean Absorbed Energy =	10.48 J (7.73 ft-lbf)
Standard Deviation =	1.07 J (0.785 ft-lbf)	Standard Deviation =	1.04 J (0.769 ft-lbf)
Range =	3.80 J (2.80 ft-lbf)	Range =	4.79 J (3.53 ft-lbf)
Coefficient of Variation =	0.068 (0.068)	Coefficient of Variation =	0.099 (0.099)

Table 2. Absorbed Energy (J)

High-Energy Specimens, HH86		Low-Energy Specimens, LL93		
Standard GROUND	EDM MODIFIED	Standard GROUND	EDM MODIFIED	EDM STANDARD
86.17	93.66	13.82	18.68	17.38
89.71	95.74	13.92	17.68	18.18
90.13	94.81	14.02	16.28	17.68
90.24	97.93	14.22	17.38	16.38
90.45	90.11	14.32	16.88	17.78
90.65	89.28	14.32	18.48	17.18
90.85	93.24	14.42	16.88	17.08
91.69	90.63	14.51	16.08	17.28
91.69	91.88	14.51	17.68	17.08
92.21	93.35	14.62	16.18	16.78
92.53	94.70	14.72	16.38	17.38
92.53	92.82	14.72	16.88	18.28
92.74	95.12	14.72	17.28	18.98
92.85	96.16	14.72	15.38	16.78
93.16	94.28	14.92	17.68	17.18
93.27	96.06	14.92	17.68	17.88
93.69	94.81	15.02	16.38	16.08
94.51	89.07	15.13	16.58	16.68
94.51	95.33	15.32	16.98	17.28
94.72	95.12	15.32	16.48	17.28
95.98	92.41	15.42	13.89	17.58
96.09	91.78	15.72	17.28	16.38
Absorbed Energy= 92.61 J Standard Deviation= 2.44 J	Absorbed Energy= 93.56 J Standard Deviation= 2.31 J	Absorbed Energy=14.78 J Standard Deviation= 0.62 J	Absorbed Energy=16.81 J Standard Deviation= 0.99 J	Absorbed Energy=17.33 J Standard Deviation= 0.65 J

Thomas A. Siewert[1] and Daniel P. Vigliotti[1]

The Effect of Charpy V-notch Striker Radii on the Absorbed Energy[2]

Reference: Siewert, T.A. and Vigliotti, D.P., "**The Effect of Charpy V-notch Striker Radii on the Absorbed Energy**," Pendulum Impact Machines: Procedures and Specimens for Verification, ASTM STP 1248, Thomas A. Siewert and A. Karl Schmieder, Eds., American Society for Testing and Materials, Philadelphia, 1995.

Abstract: The two most common Charpy V-notch striker designs (8 mm and 2 mm radii on the striking edge) were compared by using verification (reference-grade) impact specimens. Other variables in the test matrix included two different brands of U-type pendulum machines and four different specimen energy ranges (near 18, 45, 100, and 200 J). In this comparative study, we found a very small difference between the two striker designs and an even smaller difference between the two brands of machines. At 200 J, the difference between the two striker designs was about 10 J. This difference might not be important in most production testing, but must be considered in verification testing, where the acceptable range may be no more than 5 %. The standard deviations of absorbed energy for the two strikers were similar, except at 200 J, where the 2 mm striker produced standard deviations about 3 times higher than that for the 8 mm striker.

Keywords: absorbed energy, Charpy V-notch, impact test, striker radius, verification specimens

Two different striker designs are commonly found on Charpy V-notch (CVN) machines. These two common designs are described in ASTM E23 and ISO Standard R442-1965 (presently being revised by ISO TC 164/SC 4) and are distinguished primarily by the radii of the leading edge that strikes the specimen, prompting their common identification as the 8 mm and 2 mm strikers, respectively [1]. **Figure 1** compares the specified profiles near the nose of the two striker designs. The 8 mm striker is more common in the U.S. and is required for the CVN testing procedure described in ASTM Standard E 23 [2]. The 2 mm striker is more common in Europe and Asia. A few machines with the less common strikers are also found in each country, since some companies have contracts with organizations in other countries that require impact data with the other striker design. The difference in the dimensions between the two strikers is many times greater than the tolerances, so there is no possibility of producing a striker that can meet the requirements of both designs.

---

[1]Supervisory Metallurgist and Technician, respectively, Materials Reliability Division, National Institute of Standard and Technology, Boulder, Colorado 80303.

[2]Contribution of NIST; not subject to copyright.

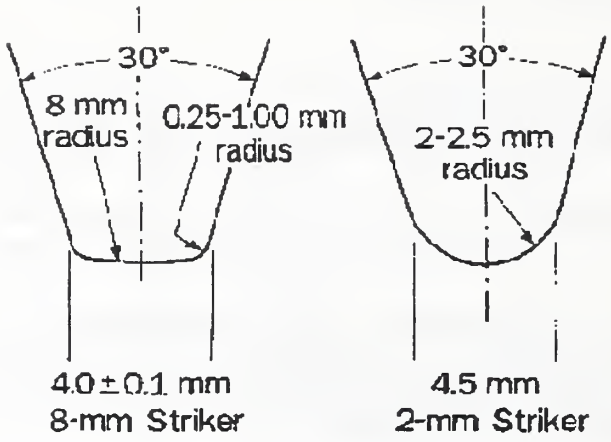


Figure 1. Comparison of the 8 and 2 mm striker dimensions.

Interchanging strikers (of the two designs) on a single impact machine is not a simple solution for obtaining data with the two strikers. One reason for not interchanging strikers is that replacing a striker can be very difficult for certain machine designs. Further, ASTM Standard E 23 requires recertification of the machine performance when the striker is changed [2]. The generally accepted justification for this requirement is that improper mounting of a striker could have such a significant effect on the machine's performance that the machine would fall outside the certification limits.

Recent comparisons of the two striker designs have reported differences in the energies. A study by Fink describes the effects of striker radius (8 mm versus 2 mm) and notching procedure on the energy [3]. He reported a linear correlation between the energies (in ft·lb) produced with the two strikers:

$$(E_{2\text{-mm}}) = 1.0420\ (E_{8\text{-mm}}) + 0.5160, \qquad (1)$$

with a coefficient of determination (r^2) of 0.9987 and a standard error of estimate of 1.36 ft·lb. In metric units, (with the number of significant digits reduced to reflect the standard error), this equation is

$$(E_{2\text{-mm}}) = 1.042\ (E_{8\text{-mm}}) + 0.70, \qquad (2)$$

where E is expressed in joules. The coefficient of determination (r^2) is 0.9987 and the standard error of estimate is 1.84 J. The study included three steels (AISI 4340, ASTM A 537, and HY-80) covering the range of 15 to 200 J.

A study by Naniwa et al. also compares the effect of the two striker designs [4]. They compared steels over a range in strengths to produce a range in CVN energies (well distributed within the range of 100 to 400 J when tested with the 8 mm striker). They reported no difference between the two striker designs for energies below 200 J, but found that the absorbed energy with the 8 mm striker was greater than that with the 2 mm striker for energies above 200 J. The difference in absorbed energies was as great as 100 J when the 2 mm striker indicated 300 J. This conclusion is in opposition to the relationship reported by Fink, who indicated that the 2 mm striker developed the higher energies.

The conflicting conclusions in these two reports indicated a need to further investigate the effects of striker design. Ideally, a relationship might be developed between the data generated with the two strikers, so that data for either striker could be calculated from the other. At least, there is a need to understand the reason for the different conclusions in these two studies. In this report, we compare the results when high-precision verification specimens were tested on two different brands of CVN machines using both striker designs.

Experimental Procedure

The primary purpose of our program was to evaluate the effects of striker design (8 mm versus 2 mm) over a range in absorbed energy, but we broadened the test matrix to include machine-specific effects. The machine-specific effects were evaluated with two brands of U-type pendulum machines.

We produced specimens with mean energies near the verification ranges currently used in the United States (18, 100, and 200 J) and additional specimens near 45 J. We used the same steel from which we manufacture the verification specimens (NIST Standard Reference Materials - SRM 2092, 2096, and 2098). We used a heat-treatable low-alloy steel for absorbed energies up to 100 J and a maraging steel for the 200 J energy. We obtained mean energies from 18 to 100 J in the low-alloy steel by varying the heat treatment (tempering temperature), with the lowest energies being produced by the treatments with the highest hardnesses. A wide range of energies was considered important because we wanted to span the ranges of the studies by Fink and Naniwa. We were most interested in the effect at 18 J because this is the energy that is most commonly the cause of a machine failing the verification test using the NIST specimens. It is also the energy closest to the requirements of many fabrication standards that require a minimum CVN energy. We do not yet have a reference-grade material with energies near 400 J, so we were unable to reevaluate the upper end of the trend noted by Naniwa et al.

Each CVN test is destructive; an individual specimen cannot be evaluated again (without complex reconstitution techniques). To allow us to compare machines, we produced our specimens in conveniently sized batches (also called series), each identified by one or two letters followed by one or two numbers. The specimens in each batch were kept together through the machining and heat treating operations to minimize the effect of processing variables. Each batch was divided into fourths at random, providing between 5 and 11 specimens from each batch for testing on each machine at each combination of energy range and striker. These data were analyzed to obtain an estimate of the standard deviation in the absorbed energy for each combination.

The use of reference-grade specimens (meeting all the requirements of E 23, but with stricter controls on their manufacturing procedures) was the most important part of our procedure. These steels and heat treatments are characterized by a narrow spread in the energy, which permits very small effects to be resolved. Verification specimens consistently have standard deviations of 5 % or less of the mean energy, which are smaller than those for most commercial steels. ASTM Research Report E28-1014 lists 2.4 J as the 95 % Repeatability Limit for these 18 J absorbed-energy verification specimens when evaluated by the ASTM E 691 interlaboratory test procedure [5,6].

We purchased two new CVN machines (from different manufacturers), each with both the 8 mm and 2 mm strikers. We selected new machines because they would reflect the latest in machine design and construction techniques. Each was of the U-pendulum design (see Ref. 2), and had a maximum capacity of 405 J (300 ft·lb). On both machines the strikers were held in place by only four bolts and could be removed and replaced easily.

Each machine was carefully mounted according to the requirements of ASTM E 23. Each was evaluated with NIST verification specimens and were certified to the requirements of E 23. We also checked each machine's performance with NIST-certified reference materials for CVN machines (5 specimens at 18 J nominal energy and 5 specimens at 100 J nominal energy) after removing and remounting the strikers. We could detect no differences in the mean or standard deviation of the specimen energies after striker replacement. Apparently these two machines have very tight tolerances for striker mounting and have mounting designs that permit accurate realignment of the striker.

To further reduce the variation in the data, a single operator performed the tests and recorded the data.. The strikers were changed between testing each series, for a total of 8 changes on each machine. A machine's performance was not checked after each change in the strikers. Our initial tests of striker replacement with NIST-certified specimens convinced us that changing the strikers on these machines had an effect smaller than we could measure. The specimens with mean energies near 18, 45 and 100 J were tested at −40 °C, and the specimens with mean energies near 200 J were tested at room temperature.

Results and Discussion

Table 1 lists the mean energies obtained for each combination of energy range, machine, and striker. **Figure 2** shows the data from Table 1 as an X-Y plot with the 8 mm data along the Y axis and the 2 mm data along the X-axis. The data fall very close to the solid line, which represents a one-to-one relationship between the two strikers, so close that it is difficult to determine the fine structure on this scale. To reveal the small differences between the two strikers, we have replotted the data in **Figure 3** as the difference in the energy (2 mm minus 8 mm) for the various combinations of the test matrix. Either the 2 mm or 8 mm data could have been selected for the horizontal scale: we arbitrarily selected the 8 mm data for this axis. **Figure 3** shows that the 2 mm striker gives higher energies than the 8 mm striker at 45 J, and vice-versa at 200 J. It also indicates that the two brands of machines have similar behavior.

Table 2 lists the standard deviation for each combination of energy range, machine, and striker, together with the number of specimens tested. We have plotted the standard deviation as a vertical bar extending one standard deviation both above and below the mean data from **Figure 3**. **Figure 4** is intended only to show how the standard deviation compares to the bias in the means as a function of energy. A detailed comparison of the standard deviations for the two strikers is included near the end of this section. **Figure 4** reveals that the standard deviation is roughly proportional to the energy, increasing with increasing energy. The means for the two different machines differ by less than one standard deviation, according to the data from **Table 2**, suggesting that the machine effect is small. However, **Figure 4** shows that the difference between the two strikers is sometimes greater than one standard deviation in the 8 mm data, at least for certain energies. The difference between the two strikers is less than one standard deviation at 18 J, slightly more than one standard deviation at 45 J, nearly 0 at 100 J, and about two standard deviations at 200 J. We interpret this to indicate that the radius of the striker nose can be an important parameter, varying in importance as a function of energy.

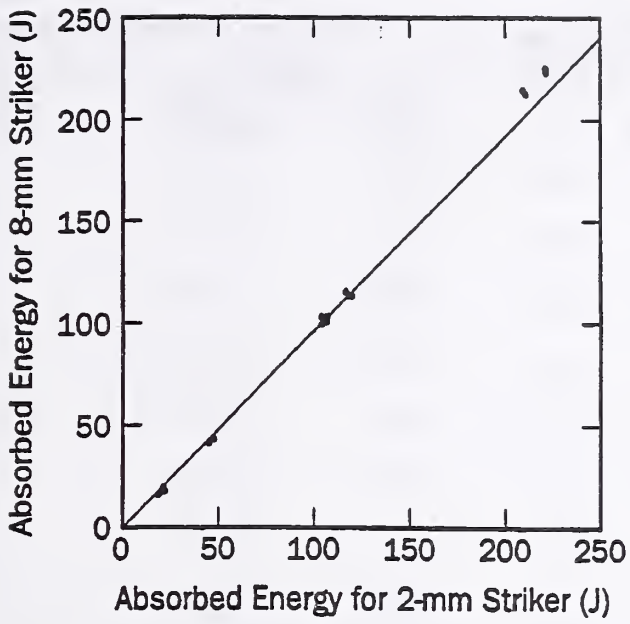


Figure 2. Energy means when tested by the two striker designs. Each point represents data from a single batch tested by the two strikers.

The data in the report by Naniwa et al were only graphical and the uncertainties in recovering the correct energies have precluded us from attempting to develop statistical measures of the scatter. Their report suggests that they performed a comparison similar to ours, splitting batches for testing by the two striker designs. **Figure 5**, which we have reproduced from their report, shows their data as the open circles on a plot with the two energies on the two axes [4]. We have added our data as solid circles, Fink's data as X's, and a line that represents our best estimate of the mean of the Naniwa data. The apparent scatter in their data suggests a standard deviation near 40 J. Unfortunately, the scale necessary to contain the Naniwa data is so coarse that the fine structure in the data below 100 J (the Fink data and our data) cannot be resolved.

To show the small differences between the two strikers, we have replotted in **Figure 6** our standard deviation data, the Fink data, and the mean for the Naniwa data in this energy range, as difference between the two strikers versus the mean. We have drawn smooth lines through both the upper and lower bounds of our standard-deviation data from Figure 4 and shaded the band between them. Although we have data gaps with our energy range that could affect the shape of the band between our data, the band shows the general trend of our data and emphasizes that the differences between the two striker designs are nonlinear. **Figure 6** indicates that the 2 mm striker (relative to the 8 mm striker) has a small positive bias that grows as the mean energy increases from 18 to 45 J, but then decreases and goes negative.

The Naniwa data appear to support the negative trend that we observed above 100 J. It is not surprising that steels with a standard deviation this large in absorbed energy would not resolve the trend noted by Fink between 100 and 200 J.

Table 1. Averages (Mean) Energy from CVN Testing of Reference-Grade Specimens

Specimen Series	CVN Machine Brand	8-mm Striker, J	2-mm Striker, J	Difference,* J
LL-39	A	18.3	18.6	0.3
	B	18.1	18.4	0.3
LL-40	A	18.5	18.8	0.3
	B	18.4	18.2	- 0.2
M-6	A	43.3	45.2	1.9
	B	44.0	45.8	1.8
HH-37	A	112.5	115.1	2.6
	B	116.7	114.0	- 2.7
HH-39	A	99.8	102.0	2.2
	B	104.7	103.2	- 1.5
HH-40	A	100.5	99.8	- 0.7
	B	100.9	101.7	0.6
SH-1	A	215.9	204.1	- 11.8
	B	216.5	200.9	- 15.6
SS-1	A	225.2	215.6	- 9.6
	B	223.8	215.3	- 8.5

* Derived by subtracting the 8 mm absorbed energy from the 2 mm absorbed energy.

Table 2. Standard Deviations of the Energy from CVN Testing of Reference-Grade Specimens

Specimen Series	Number of Specimens in Series*	CVN Machine Brand	8-mm Striker, J	2-mm Striker, J
LL-39	40	A	0.48	0.36
		B	0.60	0.35
LL-40	40	A	0.43	0.58
		B	0.86	0.44
M-6	20	A	2.25	2.02
		B	1.81	2.21
HH-37	44	A	3.61	3.85
		B	3.36	3.48
HH-39	44	A	1.81	2.88
		B	3.64	3.32
HH-40	36	A	2.77	2.29
		B	3.38	2.35
SH-1	20	A	4.70	11.60
		B	2.51	11.67
SS-1	20	A	5.41	16.56
		B	2.57	12.97

* Number of specimens in a series which was split between two machines and two strikers, so each standard deviation is based on one-fourth of this number.

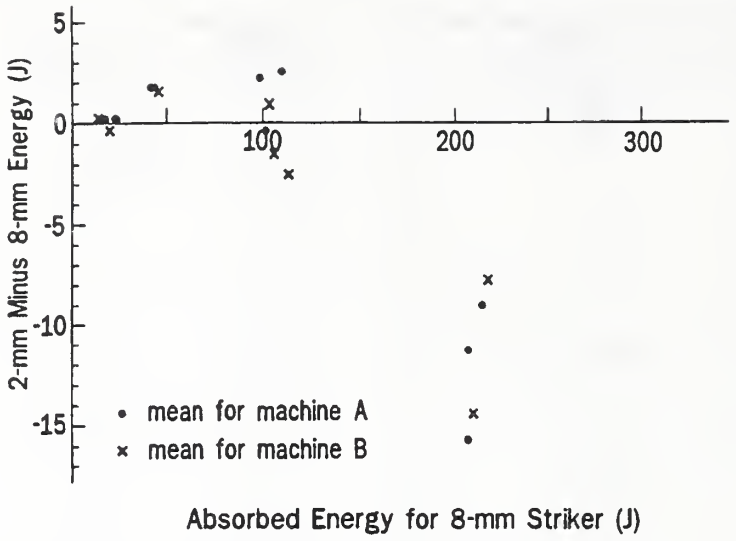


Figure 3. Graphical representation of the differences in the means for the data in Table 1 (2 mm minus 8 mm striker data for each series and each machine).

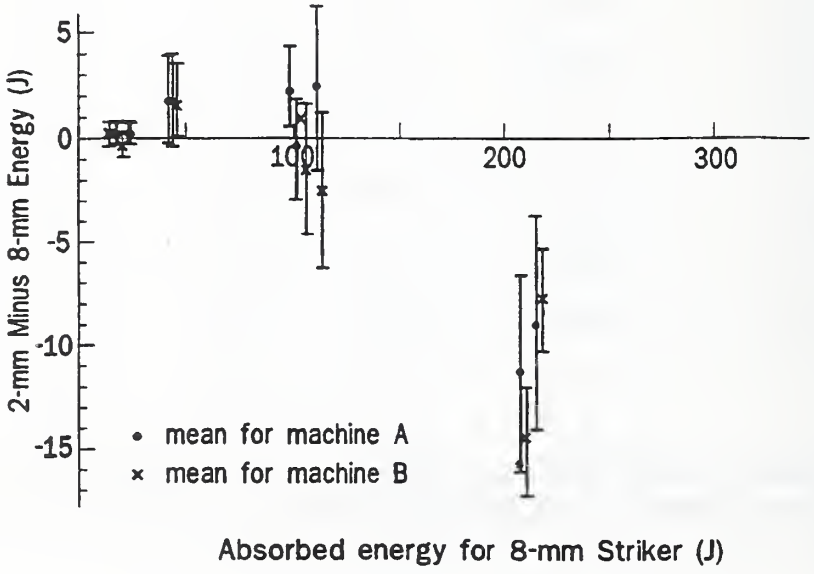


Figure 4. Graphical representation of the standard deviations for the 8-mm striker data in Tables 1 and 2.

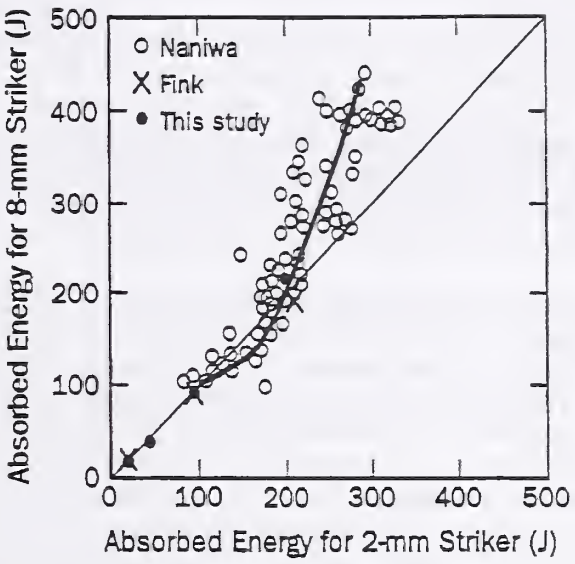


Figure 5. Data from the Naniwa study as open circles (Figure 3a in their report) to which we have added the Fink data, our data, and a line indicating the approximate mean,.

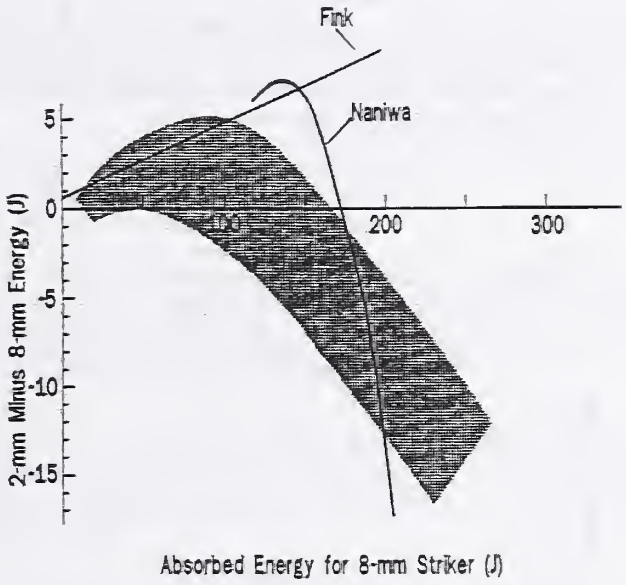


Figure 6. Comparison of our standard deviation scatter band to the mean of the data reported by Naniwa et al and Fink.

Our band follows the relationship reported by Fink [Equation (2)] up to about 100 J. In this range, Fink used materials that were very similar to ours, and this may help to explain the good fit with our data. Beyond this, our data indicate a deviate significantly negatively from his prediction. Fink's report included all his data in tabular format, so we were able to calculate standard deviations for his data [3]. Although his data near 200 J show a positive bias, the standard deviation (15 J) is sufficiently large that our mean is within one standard deviation of his mean. However, the HY-80 material that he used for his 200 J specimens has a composition and microstructure significantly different from those for the martempering steel that we used in this range, so there might be a material effect, perhaps hardness or strain hardening.

These results indicate that great caution should be used when comparing data developed on two different striker designs. Data generated with the wrong striker (a striker other than the striker specified in a standard or testing protocol) should be used only to obtain a rough estimate of material performance. Conversion of data generated with one striker to the other type of striker is subject to the uncertainty of our scatterband and additional uncertainty if the material properties are different from those used in our study. While the difference above 200 J is most dramatic, the bias between the two strikers appears to be about 0.3 J for an energy near 18 J. This bias seems small but can be important when material is very near a specification requirement. The 1.8 J bias at 45 J makes comparison of data developed with the two different strikers even more difficult in this range.

We noticed one other difference between the two designs of striker. **Figure 7** shows the standard deviation in our data by striker. Both strikers had similar standard deviations (following a linear trend) up to 100 J, but at 200 J they differed significantly. For both machines and with two different batches of specimens, we found the same result: the 2 mm striker produced standard deviations at 200 J about 3 times as great as those with the 8 mm striker. We have not yet found the reason for the difference, but it does suggest that the 8 mm striker produces more reproducible results when testing materials with absorbed energies of 200 J. Perhaps this is due to the fact that there is a greater tendency for specimens to wrap around the 2 mm striker in this energy range, rather than completely separate into two pieces.

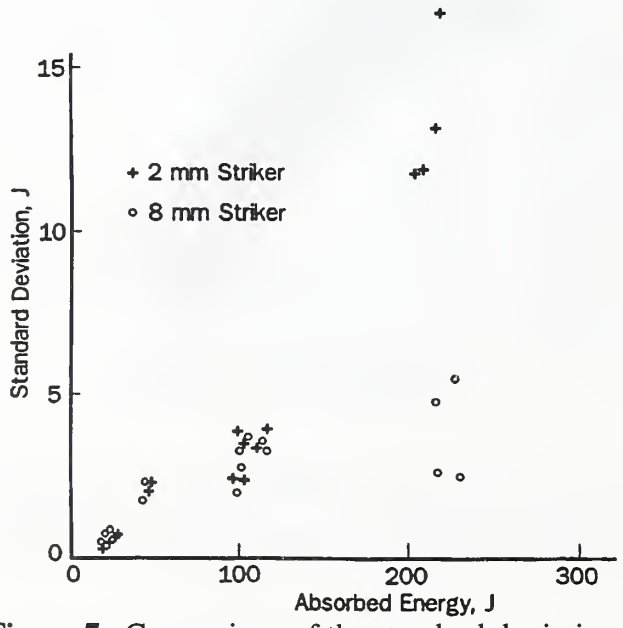


Figure 7. Comparison of the standard deviation in the energies for the two strikers.

Conclusions

1. Up to 100 J, the two striker designs produce very similar data, differing by less than one standard deviation (2 to 5 %) of the energies when measured with verification specimens. The practical effect of this difference is small, based on the qualitative nature of CVN impact testing.

2. At 200 J, the differences exceed one standard deviation, about 10 J.

3. Although the differences between the two strikers are small they must be considered in a verification program, where the acceptable range may be near 5 %.

4. The difference between energies measured using 2 mm and 8 mm strikers is complex. It is unlikely that a general relationship can be developed that will allow one machine to be certified for both strikers from a test with only one striker (except perhaps for low energies, where the difference is least).

5. Near 200 J, the 8 mm striker produces a standard deviation three times smaller than that for the 2 mm striker, but no explanation for this effect is yet apparent.

References

1. ISO Standard R442, "Verification of Pendulum Impact Testing Machines for Testing Metallic Materials," International Organization for Standards, 1965.

2. ASTM Standard E 23-92, "Standard Test Method for Notched Bar Impact Testing of Metallic Materials", Annual Book of ASTM Standards, American Society for Testing and Materials, Philadelphia, Pennsylvania, 1992.

3. D.A. Fink, "Quantitative Comparison and Evaluation of Various Notch Machining Methods and how They Affect ASTM E 23 and ISO R442 Testing Equipment Results," Charpy Impact Test: Factors and Variables. ASTM STP 1072, ed. J.M. Holt, American Society for Testing and Materials, Philadelphia, 1990.

4. T. Naniwa, M. Shibaike, M. Tanaka, H. Tani, K. Shiota, N. Hanawa, and T. Shiraishi, "Effects of the Striking Edge Radius on the Charpy Impact Test," Charpy Impact Test: Factors and Variables. ASTM STP 1072, ed. J.M. Holt, American Society for Testing and Materials, Philadelphia, 1990.

5. Research Report: E28-1014, "Interlaboratory Test Study for the Determination of Precision and Bias in Charpy V-notch Impact Testing," D.A. Shepherd and T.A. Siewert, available from American Society for Testing and Materials, Philadelphia, 1991.

6. ASTM Standard E 691-87, "Practice for Conducting an Interlaboratory Test Program to Determine the Precision of Test Methods," Annual Book of ASTM Standards, American Society for Testing and Materials, Philadelphia, 1987.

Christopher N. McCowan[1], James W. Dally[2], Daniel P. Vigliotti[1] and Ouk S. Lee[3]

Low Cost Lower Bound Toughness Measurements

Reference: McCowan, C.N., Dally, J.W., Vigliotti, D.P., Lee, O.S., "Low Cost Lower Bound Toughness Measurements," Pendulum Impact Machines: Procedures and Specimens for Verification, ASTM STP 1248, Thomas A. Siewert and A. Karl Schmieder, Eds., American Society for Testing and Materials, Philadelphia, 1995.

Abstract: A method is proposed for determining the lower bound toughness of engineering steels based on impact loading of a modified Charpy specimen in a standard Charpy machine. The modified Charpy specimen employed here is nearly the same as the over-sized specimen studied previously by Bonenberger et al. The primary difference is that a sharp notch is employed to initiate the crack instead of pre-compression. A master curve is developed to relate the stress intensity developed at the tip of the notch to the time of crack initiation. An instrumented tup is then employed to give a force-time trace that can be interpreted to estimate the crack initiation time. Results are presented for the lower bound toughness of ASTM A 533, Type B, Class 1 steel over the temperature range from 0 to 40°C. These results are higher than the results due to Bonenberger et al. for the same steel. The differences are attributed to the different techniques used to develop the starter crack in the modified Charpy specimen.

Keywords: Charpy V-Notch, Instrumented Impact Test, Lower bound toughness

During the past 35 years, significant progress has been made in developing the theory of fracture mechanics, and in perfecting test methods for measuring crack-initiation toughness. Nevertheless, failures of engineering structures continue to occur with serious consequences. Many of these failures are due to the fact that the material employed in the fabrication of the structure was not tested or certified prior to its use in construction [1]. In many instances, handbook values of fracture toughness were used in the fracture-mechanics analysis. This practice is very dangerous because the fracture toughness varies markedly from heat to heat of steel, and the values quoted in handbooks for a particular alloy should be considered as only typical.

It is necessary to test the material specified in the design to establish the crack-initiation toughness over the temperature range expected in the service of the structure. Test methods used to determine the toughness, ASTM E399 [2] for example, are well known. What is less well known to the engineer performing the analysis, and to the designer of the structure, is the huge amount of scatter encountered in measuring the initiation toughness. Recent results by Link and Joyce [3] from an extensive series of tests with A 533, Type B reactor-grade steel, presented in **Fig. 1**, illustrate the amount of scatter for steels commonly employed in pressure vessels.

[1] Materials Reliability Division, NIST, Boulder Laboratories, Boulder, CO 80303
[2] Mechanical Engineering Department, University of Maryland, College Park, MD 20742
[3] Mechanical Engineering Department, University of Inchon, Inchon, Korea

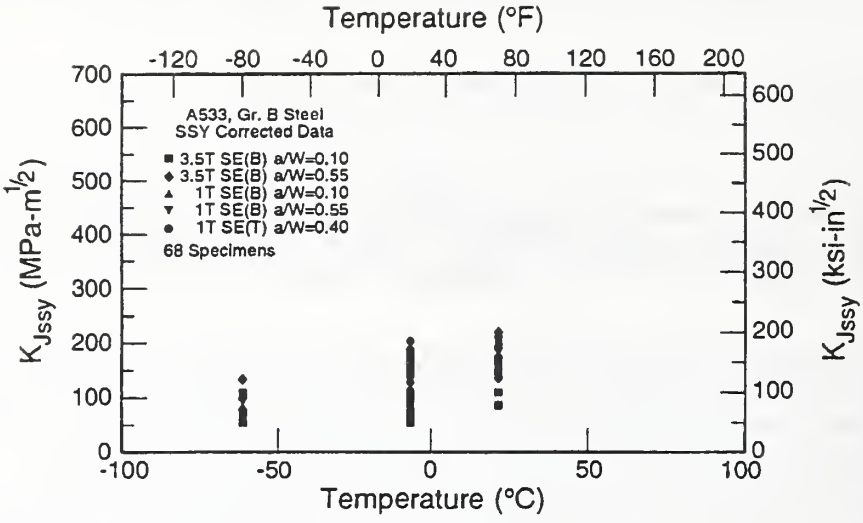


Figure 1: The scatter in the crack initiation toughness of A 533 Type B reactor grade steel is very large even near the nil ductility temperature. Data from Link and Joyce[3].

The scatter is so large (400% at the nil ductility temperature) that the only conservative method of design is to determine the lower bound, which is a curve of toughness as a function of temperature drawn below all of the test data. To establish this lower bound toughness, it is necessary to conduct many tests at each temperature to ensure that at least one or two "low" values are represented in the data set. It appears that the difficulty in measuring the lower bound, and the excessive time and expense involved in testing, are important factors that limit the effectiveness of design against fracture.

This paper describes a low-cost method of measuring the lower-bound toughness of steels that has potential to be used to certify steels for a wide variety of structural applications. The method is based on testing an oversized Charpy specimen in a standard 240 J Charpy impact machine. The Charpy specimen, shown in **Fig. 2**, is oversized in cross section with a height of 19 mm and a thickness of 12.7 mm, but is of standard length. Side grooves are cut into the specimen 2.0 mm deep with a sharp tip cutter having an included angle of 45°. The face groove, opposite the impact point, is also 2 mm deep, as in the standard Charpy specimen, but the radius of the tip is approximately 0.06 mm. This approach is similar in many respects to the method described previously by Bonenberger et al. [4,5]; however, two important modifications have been made to improve the method.

First, in preparing the modified Charpy specimens Bonenberger et al subjected the specimens to axial pre-compression stresses well above the yield strength of the material. The purpose of the pre-compression was to deform the material in the local neighborhood of the notch, thereby sharpening its tip. The practice of pre-compression may be

objectionable, because the material local to the notch is strain-hardened and its toughness may be degraded. In the method described in this paper, pre-compression is not necessary because cleavage may be initiated at temperatures in the lower transition region, if the machined notch is sufficiently sharp.

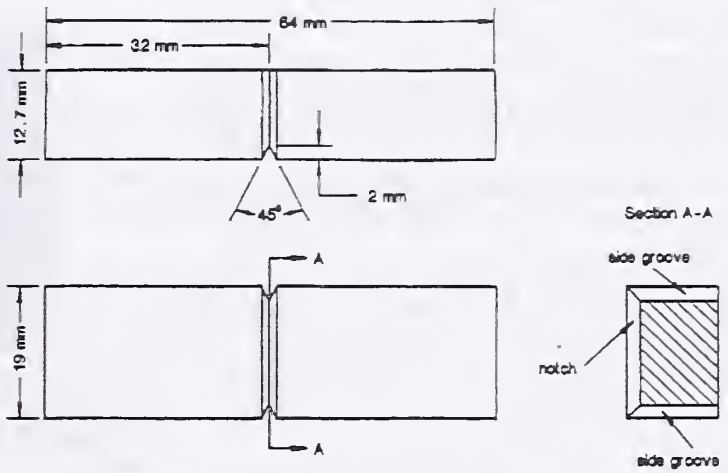


Figure 2: Geometry of the modified (oversized) Charpy specimen.

Second, Bonenberger et al employed four strain gages on each specimen that were located near the crack tip to record a strain relative to time during the impact period. The strain-time trace was interpreted to give the stress intensity factor as a function of time. The time of crack initiation was evident from the strain-time traces because of the rapid decrease in the strain at the gage location due to crack extension. The value of the stress intensity factor K_I, at the initiation time, was taken as K_{Id}, which is an estimate of the lower-bound toughness. The approach described in this paper involved strain gaging and testing several modified Charpy specimens to develop a master curve showing the stress intensity factor K_I as a function of time. Strain gaging can then be discontinued for subsequent evaluations of the lower-bound toughness. An instrumented tup is employed to establish the time of the initiation of the crack at the notch. Using the initiation time together with the (K_I-t) master curve, it is possible to measure the lower-bound toughness without strain gaging at test temperatures in the lower transition region.

Eliminating the need for pre-compression removes the question of altering the material properties prior to testing. Eliminating the strain gages reduces the time and cost of lower-bound toughness measurements by at least a factor of ten.

SPECIMEN PREPARATION

The modified Charpy specimens were machined from A 533 B reactor grade steel. This particular lot of material was available from the round-robin test series conducted to verify the arrest toughness testing procedure, ASTM E-1221 [2]. The properties of the material are well known and are described in reference [6]. It is important to note that the RT_{NDT} was -2°C, because it indicates the temperature range of interest in defining the lower-bound toughness in the lower transition region. At room temperature this plate of A 533B exhibited a yield strength of 480 MPa, and an arrest toughness K_{Ia} of 83.4 $Mpa{\cdot}m^{1/2}$.

The specimens were machined in accordance with the drawing shown in **Fig. 2**. Of particular concern was the sharpness of the notch, since it was to act as a starter crack. The notch was machined with a multi-tooth carbide cutter with a 45° included angle. The tip of the cutter ground with a tip radius of 0.04 mm; however, we expected that tool wear would increase that radius. We checked the profile of several notches using shadow projection with high magnification and found some variation in notch quality, with notch

radii varying from 0.05 to 0.07 mm. It was also noted that the notch tip was not a perfect circular arc. A typical profile of a notch tip, presented in **Fig. 3**, shows the quality typical of the geometry maintained at the notch tip.

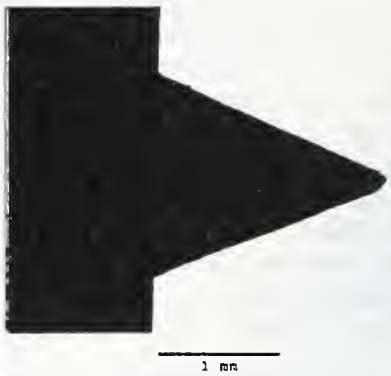


Figure 3: Micrograph of a replica of the notch.

Four bonded-foil strain gages, with an active element 1.5 mm long, were installed on each specimen 5 mm from the tip of the notch, as indicated in **Fig. 4**. This location is identical to that employed by Bonenberger et al [5] in the initial development of this approach. During the impact experiments, the strain gage signals were recorded on a four-channel digital oscilloscope. The frequency response of the strain measuring system was controlled by the strain-gage amplifiers. Whose gains were essentially flat from dc to 100 kHz.

CHARPY MACHINE MODIFICATIONS

The Charpy machine employed in the study had a standard U type pendulum with a capacity of 240 J, although we made two modifications to the machine. First, to accommodate the added height of the oversize specimen (19 mm versus 10 mm for the standard specimen), the rails supporting the specimen were reduced in thickness by 4.5 mm. This change maintained the position of the centerline of the specimen relative to the nose piece on the hammer.

The second change involved instrumenting the nose piece, to permit a signal representing the impact force to be recorded. Shallow slots were ground on both sides of the nose piece of the hammer to accommodate semiconductor strain gages. Cover plates were bonded to the nose piece to protect the gages from contact with the Charpy specimens during the fracture process.
The sensitivity of the instrumented nose piece was very dependent on the location of the applied force along its height. We resolved this problem by an in situ calibration technique based on absorbed energy. We first calibrated the nose piece in a testing machine to obtain an approximation of the calibration constant for its load sensitivity. Subsequently, this first estimate was adjusted so that the absorbed energy as calculated from the force-time record coincided with the absorbed energy measured directly from the Charpy machine. The repeatability of this calibration technique was excellent.

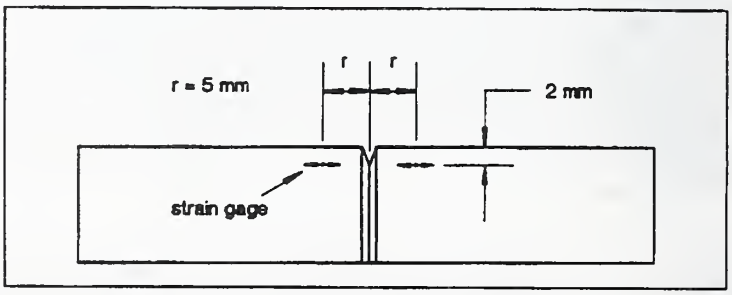


Figure 4: Location of the strain gages relative to the notch tip.

DEVELOPING THE MASTER CURVE

A typical set of strain-time traces recorded from the four gages is shown in **Fig. 5**. We note that each gage indicates a slightly different signal depending on the response of specimen. To accommodate these differences, the signals from the four different gages are averaged together to give a single strain-time trace as indicated in Fig. 6.

The strain-time trace in **Fig. 6** clearly indicates the dynamics of the modified Charpy specimen. The strain increases with time in a nearly linearly manner for the first 50 μs, and then becomes more nonlinear. The nonlinearity is due to two effects. First, with increasing time the Hertzian load at the contact point produces local compressive stresses that exceed the yield strength of the A 533B steel, and a portion of the hammer displacement is accommodated by local plastic deformation at the contact point. Second, with the higher strains (1200 με and above) the regions near the notch tip have yielded, and the specimen deformation is accommodated by plastic deformation near the notch tip. It is evident that the non-linearity in the strain-time trace is to be expected, and that it will be dependent on the yield strength of the material tested and the shape of the nose piece on the Charpy machine.

Oscillations are also evident in the strain-time record presented in **Fig. 6**. The first oscillation has a period of 40 μs and the second has a period of about 65 μs. The third oscillation begins, but is damped beyond recognition and its period is not evident. Clearly, the specimen is undergoing vibration during impact; however, the magnitude of the oscillations is small compared to the mean strain later (after 70 μs) in the impact event. Since all of the specimens tested in this series of experiments failed at times longer than 95 μs, it appears that the effect of specimen vibration on the strain recorded at the time of crack initiation was relatively small.

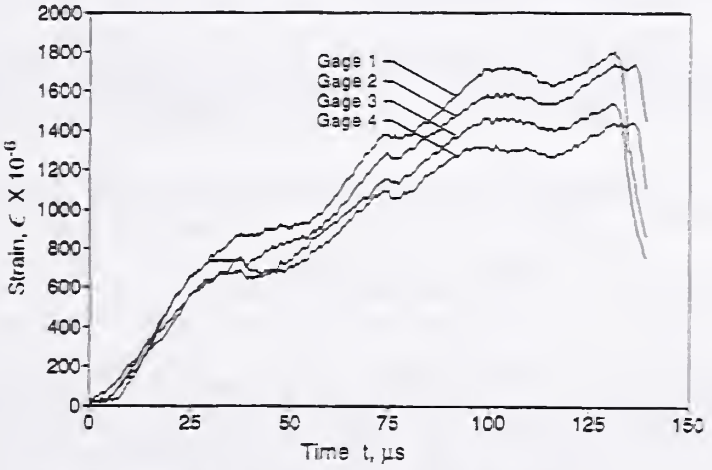


Figure 5: Typical strain-time traces showing gage-to-gage variation in strain.

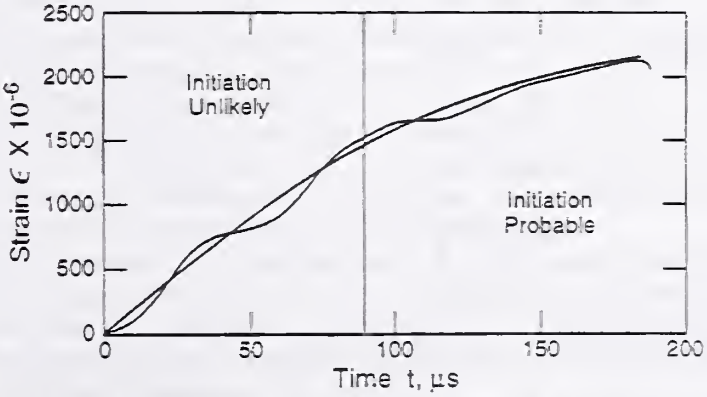


Figure 6: Dynamic characteristics of the strain-time trace.

The reproducibility of the strain-time traces is illustrated in **Fig. 7**, where the records for four different specimens tested at temperatures varying from 0 to 40 °C are presented. Examination of these results show that the strain-time traces are almost identical for times less than about 75 μs and strains less than 0.0014. Later in the impact event larger record-to-record deviations are evident. It is believed that these deviations are due to the tearing that occurs at the notch tip before cleavage failure is initiated. In the section on Fractography, we will show evidence of this tearing and describe the specimen-to-specimen variation of the extent of tearing observed at the tip of the notch.

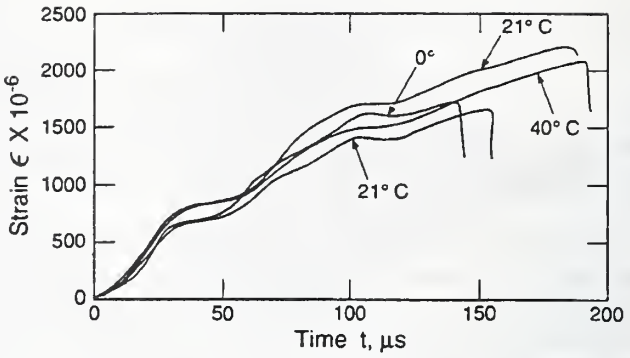


Figure 7: Four stress-time traces showing reproducibility.

The master strain-time curve, shown in **Fig. 8**, is representative of the individual traces presented in **Fig. 7**. It is necessary to convert this master strain-time curve to a corresponding master curve for stress intensity factor versus time.

This conversion was made by employing the relationship derived by Bonenberger et al. [5]:

$$K_{Id} = 38,880\epsilon^*_{yy} \qquad (1)$$

where K_{Id} has units of MPa·m$^{1/2}$, and ϵ^*_{yy} is the strain at the instant of cleavage initiation. We consider the dynamic initiation toughness K_{Id} as well as the crack arrest toughness K_{Ia} to be good estimates of the lower-bound toughness. Accordingly the value of K_{Id} is presented on the right hand ordinate of **Fig. 8**. The smooth nonlinear shape of the master curve suggested that it would be possible to fit the curve with a relatively simple relation. We have selected a two-term function of the form shown below to relate toughness with time.

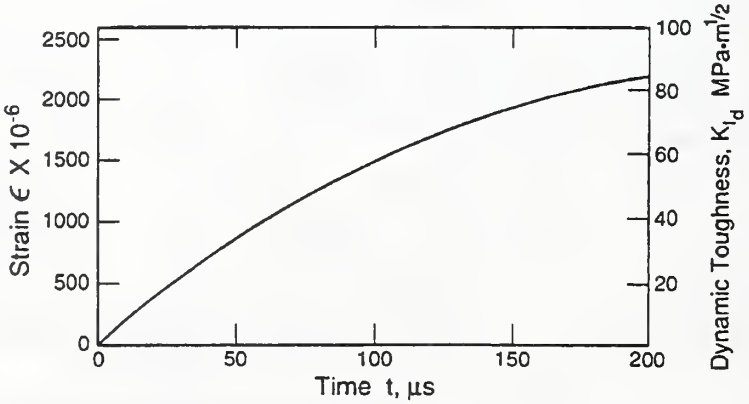


Figure 8: Master curve for ϵ-t and K_{Id}-t

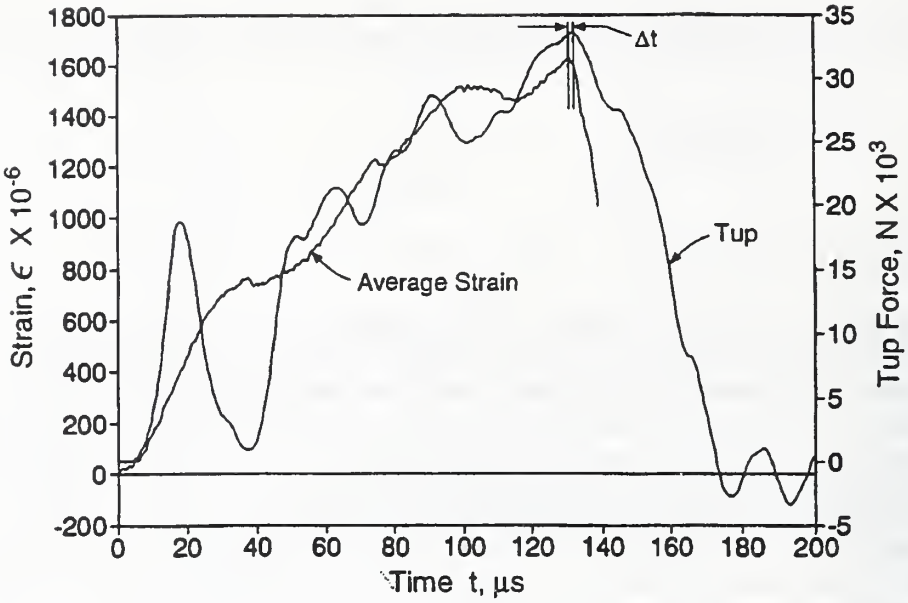
In Eq. (2) the time t is in μs and K_{Id} is in MPa·$m^{1/2}$. The constants a and b were determined as 4.88 and 0.0758 by fitting the relation to data taken at t = 100 and 200 μs.

$$K_{Id} = at^{1/2} + bt \qquad (2)$$

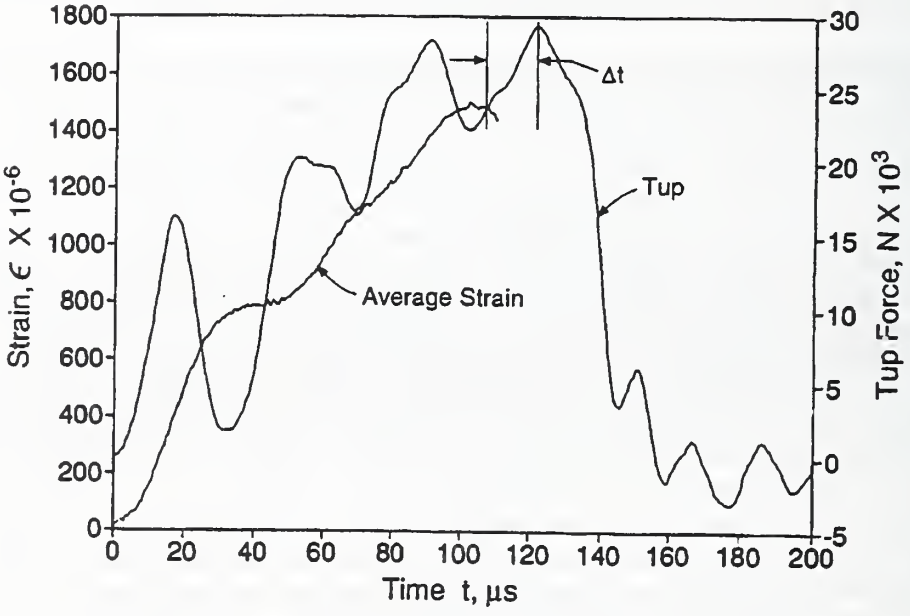
CRACK INITIATION TIME

The dynamic initiation toughness for A 533B reactor grade steel may be determined from the master curve shown in **Fig. 8** or from Eq. (2), if some method is used to measure the time at which the crack initiates during the impact event. There are three different methods for measuring the initiation time. Strain gages can be mounted near the crack tip, and their signals will sharply decrease within a few microseconds after the crack initiates [4,5]. A second technique involves the placement of a coil near the Charpy specimen to sense the magnetic field. When the crack initiates the field suddenly changes. Recording the voltage produced by the coil with respect to time gives a discontinuity in the voltage-time trace that indicates the initiation time [6]. The third technique involves measuring the impact force as a function of time during the impact event. The tup signal decreases rapidly with time when the crack initiates (in cleavage) and the specimen stiffness shows a corresponding decrease. We employed the both the tup and strain gage signals in determining initiation time in this study.

Oscilloscope records of the average strain-time and tup force-time traces are shown in **Fig. 9**. We include four records to show the effect of different testing temperatures. An examination of these records indicates that the tup signal oscillates with very large amplitude for the first cycle; however, in subsequent cycles the amplitude of the oscillation decreases markedly but does not vanish. The oscillations in the force-time records are due to several different harmonics. The fundamental harmonic, at 27 kHz, is probably due to the ringing of the nose piece at its natural frequency. The natural frequency of the specimen (about 2.5 kHz) is not evident in the tup record. For times larger than about 50 to 60 μs, the tup signal appears to track the strain signal with oscillations producing the primary deviations between the two signals. When the crack initiated in the modified Charpy specimen, both signal decreased rapidly with respect to time. The difference in the indicated time of initiation varied from specimen to specimen. Reference to **Fig. 9a** shows a very close correspondence, because the oscillation of the tup signal was in phase with the initiation of the crack. That is, the oscillation produced a decrease in the force signal at the same instant that the reduced stiffness of a failing Charpy specimen produced a decrease in the tup signal. On the other hand, **Fig. 9b, c, and d** show that the oscillation was producing an apparent increase in the force signal, at the instant of crack initiation. In this case, the force-time trace is delayed in its response to the reduced stiffness of the specimen after crack initiation. It is clear then that differences occur in the estimate of the crack initiation time as measured from the strain-time traces and the force-time traces. The differences are indicated by the graph of initiation time as measured from the strain-time trace on the ordinate, and initiation time as measured from the force-time traces on the abscissa in **Fig. 10**. For ideal measurements, the initiation times associated with both methods should be equal and the data points should track the no-delay line shown **Fig. 10**. However, neither measurement technique is without error, and we observe a dispersion of data points on both sides of the no-delay line. Most of the points indicate that the indications from the force-time records are delayed relative to the time indicated from the strain-time trace.

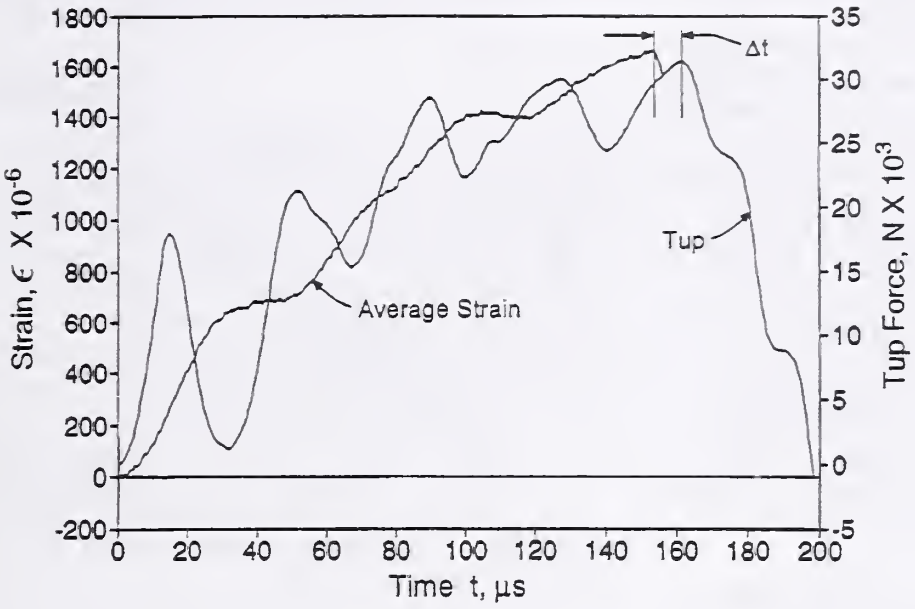


(a)

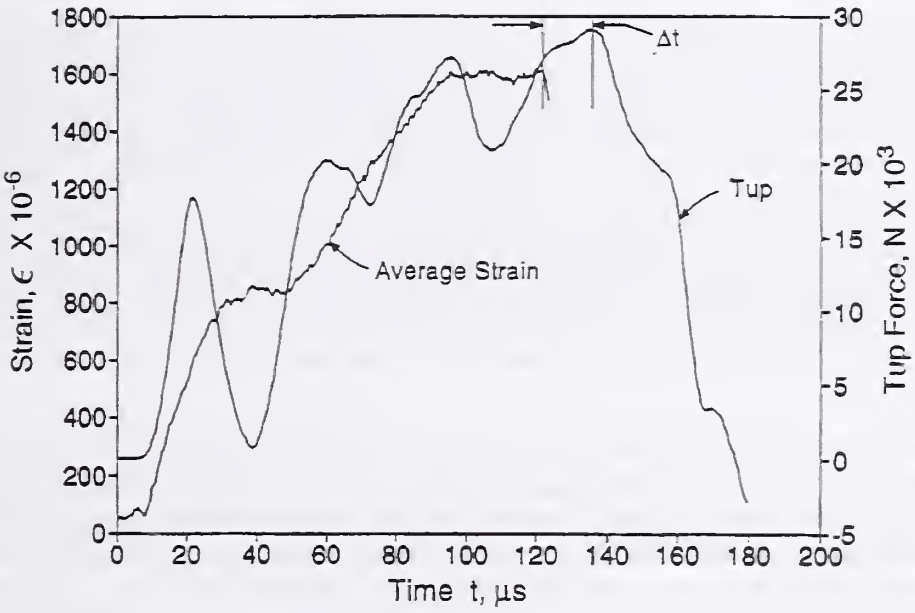


(b)

Figure 9: Four records showing correspondence between the strain gage signals (average) and the tup signal: (a) E3-11 at 0 °C, (b) E3-3 at 10 °C; (c) & (d), next page.



(c)



(d)

Figure 9 continued: (c) E3-16 at 21 °C, and (d) E3-17 at 30 °C.

The delay in most cases is less than 10 μs. In three cases the force time record gives an early indication of initiation, but the differences in these early signals is usually less than 5 μs. We note three data points in **Fig. 10** that fall well outside the ±10 μs band. In two instances the specimen yielded, and the crack experienced a significant degree of tearing before cleavage initiated at times greater than 250 μs. The effects of tearing are to delay the fracture event and to produce an apparent toughness well above the lower-bound toughness. In one case, Q3-21, a reasonable toughness was predicted with crack initiation time of 221 μs. In general one might disregard any test with a crack initiation time significantly greater than about 200 μs.

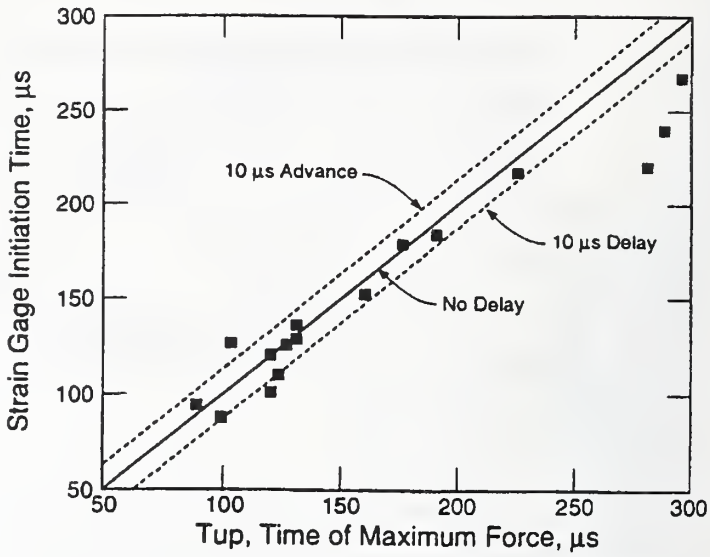


Figure 10: Comparison of initiation time determined from strain gage and tup signals.

The effect of small errors in measuring the crack initiation time depends on the time of crack initiation. If the dynamic crack initiation toughness is relatively low (about 60 MPa·m$^{1/2}$), the specimen will fail early in the impact event, say at 100 μs. A delay error of 10 μs (10 % in this instance) produces an error in K_{Id} of 6.0 %. The mitigation of the error is due to the nonlinearity of the master curve shown in **Fig. 8**. At higher toughness values (say 73 MPa·m$^{1/2}$ which corresponds to an initiation time of 150 μs) a delay error of 10 μs (6.7 % in this case) produces an error in K_{Id} of only 4 %. For the highest toughness, about 85 MPa·m$^{1/2}$, the time to initiation is 200 μs, and the error in the time measurement will be 5 % and in the toughness measurement only about 3 %. When compared to the scatter of several hundred per cent in toughness measurements typically observed in extensive and carefully controlled testing programs for K_{Ic}, these errors are very small.

We conclude that force-time records from instrumented Charpy machine can be interpreted to accurately estimate the initiation time. The estimate will usually give an initiation time that is too long by 5 to 10 μs, but this error is small compared to the typical scatter in the measurement, and the error is mitigated in part by the nonlinearity of the master K_{Id}-t curve.

RESULTS

A total of 16 modified Charpy specimens were tested in this study. Each specimen was instrumented with strain gages, and simultaneous records of strain and force were made as a function of time during the impact event. Specimens were tested at temperatures of 0, 10, 21, 30 and 40 °C. The results from the test series include the crack initiation time measured from the strain gage and instrumented tup records; the dynamic crack initiation toughness K_{Id} and the energy absorbed are shown in **Table 1**. Due to excessive tearing prior to cleavage initiation, two specimens, E3-8 and -16, were

considered to be invalid tests. The results for the crack-initiation toughness of the remaining specimens are also presented in **Fig. 11**.

FRACTOGRAPHY

The fracture surfaces from five different specimens were examined in a scanning electron microscope to study the initiation of fracture from the "sharp" notch tip. The results that were observed varied considerably from specimen to specimen. The smallest amount of tearing occurred in specimen E3-11, as illustrated in **Fig. 12**. An overall view of the notch region at 20X, presented in **Fig. 12a**, shows a narrow tearing zone at the notch tip prior to initiation of cleavage over a large region. It is important to note that the extent of tearing is not uniform, but the depth of the tear zone varies over the height of the specimen. This fact is better illustrated in **Fig. 12 b**, where the transition from ductile tearing to cleavage is presented at a magnification of 700X. We observe that the depth of the tear zone varies from about 33 μm at the bottom of the fractograph to about 47 μm at the top. The cleavage region shows small areas of ductile hole joining at the ridge lines, which is typical for this type of steel when cleavage occurs at test temperatures close to the RT_{NDT}.

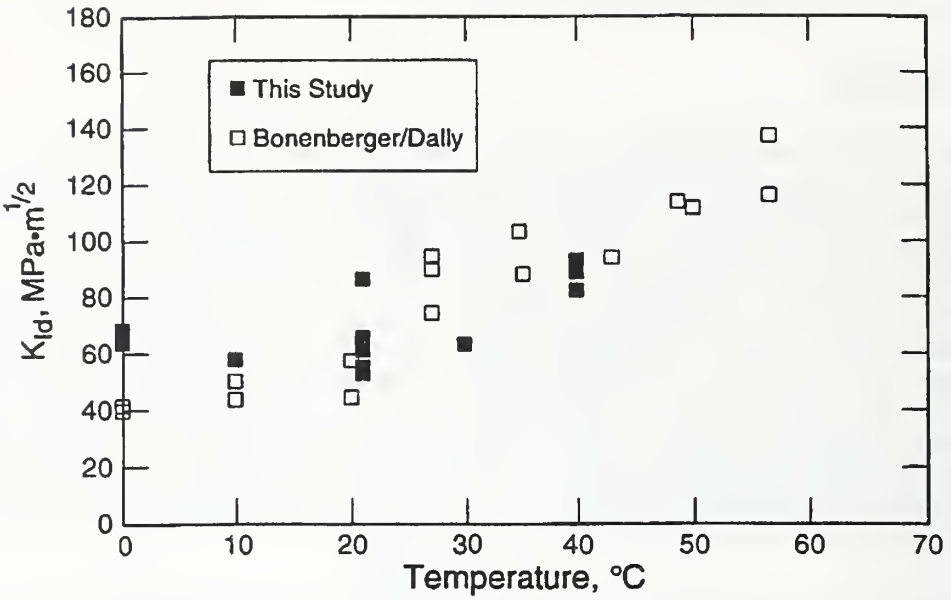
The most extensive tearing occurred for specimen E3-8 which was tested at 30°C. The fractograph shown in **Fig. 13a** indicates that the depth of the tear zone exceeded 1 mm. The propagation of the crack front by the tearing mechanism is relatively slow in comparison to cleavage propagation. As a consequence, the time required for cleavage initiation is very long. In this case the time to initiation was 267 μs. With this large amount of tearing the test was considered invalid and the data was not used to determine the lower-bound toughness. The transition from tearing to cleavage eventually occurs, as shown in **Fig. 13 b**. Again the cleavage region shows evidence of late-breaking ligaments at the ridge lines of the cleavage facets. The ductile hole joining in these regions is typical of this steel when tested at 30 to 40 °C above the RT_{NDT}.

The area of tearing at the tip of the notch varied from 1.8 mm^2 to 9.6 mm^2. The time to crack initiation was a function of this area, as expected. The shorter initiation times (less than 150 μs) were associated with tearing areas of less than 2 mm^2, and the longer initiation times (more than 200 μs) were associated with areas in excess of 2.5 mm^2.

DISCUSSION AND CONCLUSIONS

The concept of determining crack initiation toughness from Charpy specimens is not new. Initial studies [8,9] coincided with the development of instrumented tups, and efforts were made to relate the toughness directly to the impact force. The method for determining K_{Id} from the tup records was described by Radon and Turner [10], and this method was employed by Server and Tetelman [11] in their studies of a reactor grade steel. Both of these investigations used standard Charpy specimens that had been precracked in fatigue and side-grooved. From these studies, guidelines [12] were developed to determine K_{Id} and $_{Jid}$.

The difficulty in using the standard Charpy specimen to determine toughness is that the small height (10 mm) of the specimen is sufficient to provide plain strain constraint for only those materials with extremely low toughness. For more commonly employed engineering steels, additional constraint is necessary to accommodate increased



(a)

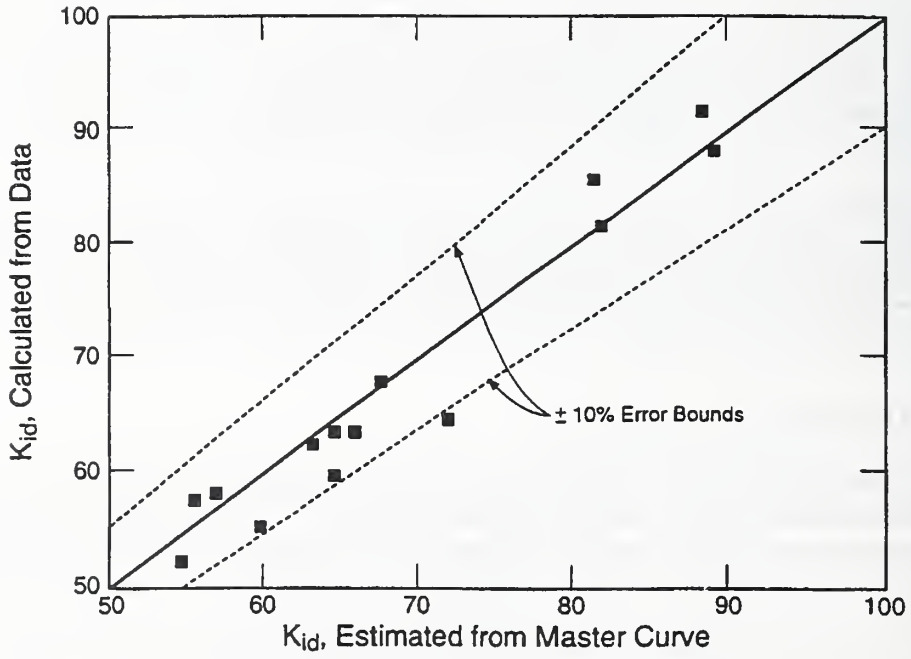
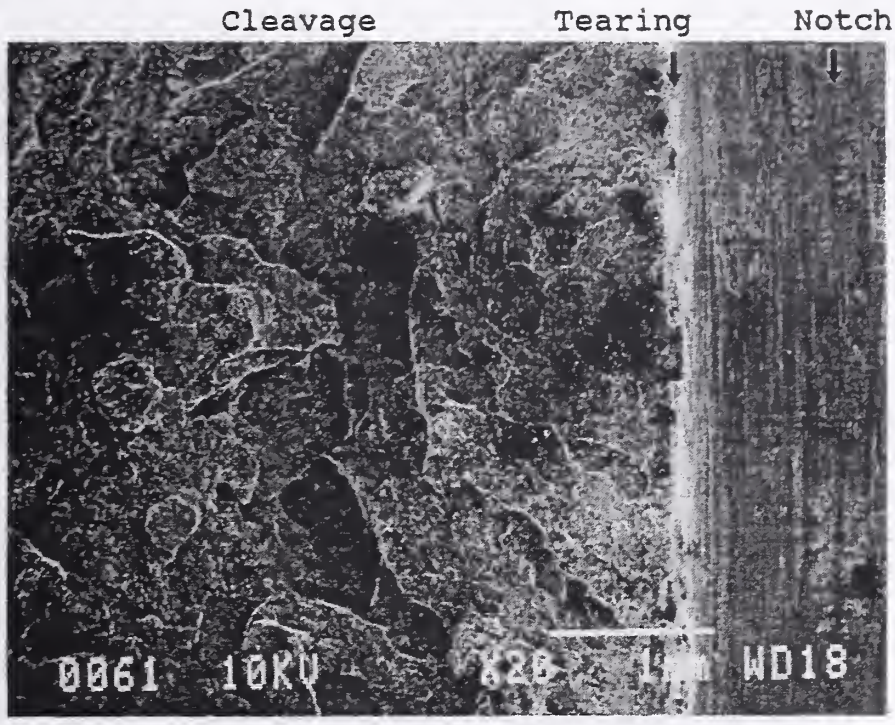
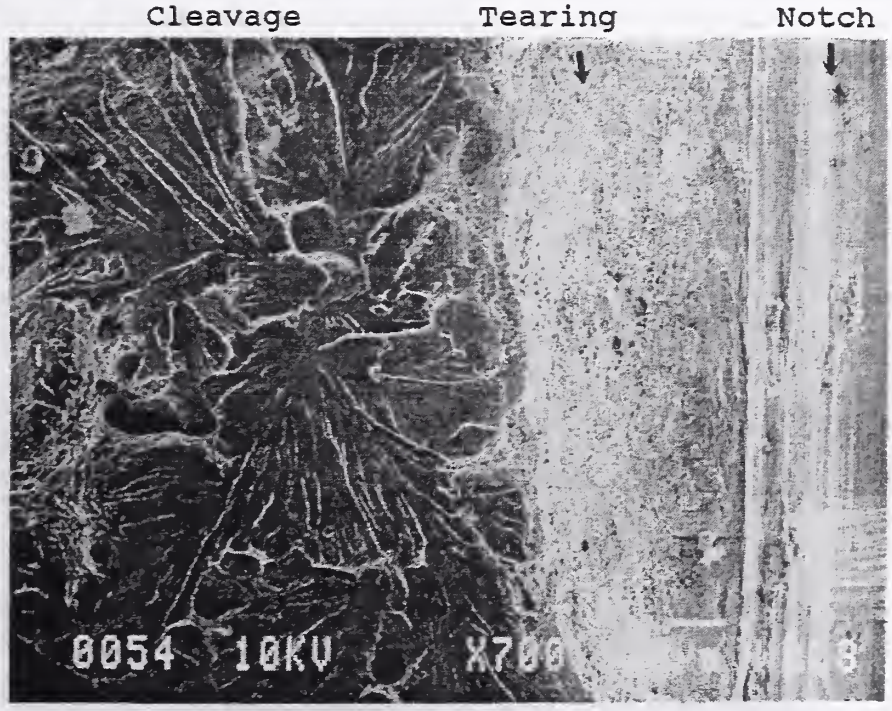


Figure 11: (a) The K_{Id} temperature relation for A553 B reactor-grade steel, and (b) the correlation between K_{Id} values calculated from the data with K_{Id} values estimated from the master curve.

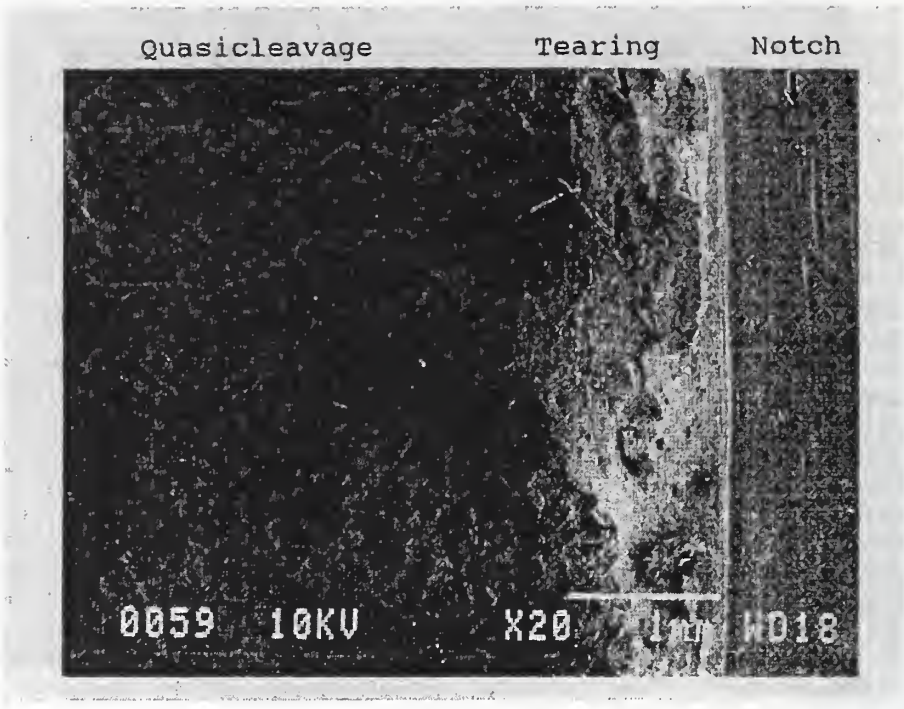


(a)

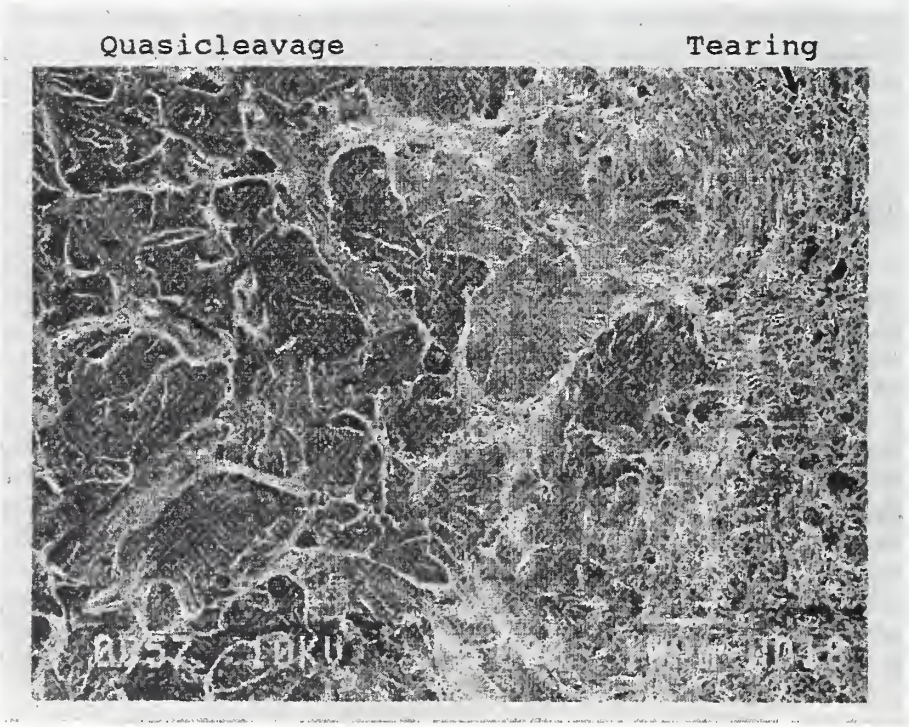


(b)

Figure 12: The tearing zone at the root of the notch in the E3-11sample (0 °C), is approximately 0.1 mm wide. Tearing was followed by initiation of cleavage, which is shown more clearly at higher magnification, b.



(a)



(b)

Figure 13: The tearing zone at the root of the notch in the E3-8 sample (30 °C), is over 1 mm wide in some regions. Tearing was followed by initiation of cleavage (mixed mode), which is more clearly shown at higher magnification, b.

TABLE 1--Results from the Modified Charpy Test Series

Spec. No.	Temp. °C	K_{Id} $MPam^{1/2}$	Strain Time μs	Force Time μs	Absorbed Energy J
E3-11	0	70.4	131	132	32.2
Q3-20	0	75.6	137	132	32.9
E3-3	10	71.2	102	121	44.8
E3-20	10	63.9	98	100	26.6
E3-1	21	60.4	127	139	29.4
E3-6	21	58.2	95	95	23.1
E3-10	21	61.5	111	124	35.0
E3-16	21	71.7	154	161	37.8
E3-18	21	70.4	127	128	28.0
Q3-17	21	95.4	131	132	55.3
E3-7	30	69.0	180	177	35.0
E3-17	30	69.3	122	121	29.4
E3-8	30	100.1	267	296	100.8
Q3-10	40	101.5	218	227	59.5
Q3-21	40	99.0	221	283	79.8
Q3-23	40	90.6	191	*	56.0

* Instrument malfunction on tup record.

toughness. Hoyt [13] had suggested using double-height specimens to increase constraint as early as 1938, but nothing was done to implement this suggestion until Bonenberger developed the modified specimen geometry [4] with a height of 19.2 mm. This specimen offers sufficient constraint to permit testing of engineering steels with K_{Id} approaching about 100 Mpa·$m^{1/2}$.

We have extended the work of Bonenberger et al. by modifying the procedures followed for specimen preparation and testing. We have employed a sharp notch instead of a precrack formed with precompression. The results obtained indicate that cleavage can be initiated from the sharp notch, but that the cleavage is preceded by a small amount of tearing. The depth of the tearing ranged from 30 to 50 μm for specimens exhibiting lower toughness (55 to 70 MPa·$m^{1/2}$) to as much as 1 mm for specimens with very high toughness (in excess of 100 MPa·$m^{1/2}$). The tearing transitions into cleavage even in the high-toughness specimens; however, tearing is a slow process and the initiation time is extended to 200 μs or longer.

We have also shown that an instrumented tup can be employed to estimate the crack initiation time. There is usually a small delay in the response of the tup, when compared to the strain-gage response; however, the error produced by the typical delay (5 to 10 μs) is negligible when compared to the scatter observed in carefully controlled and standardized tests for toughness.

A master curve relating crack initiation time was developed in this investigation. This curve differs from the master curve developed by Bonenberger et al [5], even though the material was from the same plate. The curve presented in **Fig. 8** is nearly identical to the curve in [5] for small strains, but for the larger strains the curve developed in this study shows more nonlinearity. We believe that part of the differences may be due to the different Charpy machines that were employed. Bonenberger employed a U-type machine with a capacity of 400 J. The machine was old (worn tup), used extensively by undergraduate students, and not firmly mounted to the floor. The machine used in this study was a U-type with a capacity of 250 J. It was in excellent condition and grouted and properly bolted to the floor with a suitable foundation. The nose piece on the hammer was new. We believe that a significant part of the nonlinearity in the master curve is due to the plastic deformation at the contact point due to Hertzian stresses that exceed the yield strength.

A second difference was in the material at the notch. Although all the specimens were from the same plate, Bonenberger pre-compressed the material in the local neighborhood of the notch. This pre-compression elevated the local yield strength and the specimen exhibited better resistance to the Hertzian contact stresses and required a longer time before yielding at the net section occurred. The precompression also introduced residual tensile stresses that were partially relieved by short cracks at the notch tip.

The third difference was in the depth of the side grooves. In this study we employed a depth of 2 mm compared to a depth of 1.9 mm used by Bonenberger et al. This difference would tend to cause yielding of the more deeply side-grooved specimens at an earlier time and to accentuate the non linearities.

The results that we obtained for K_{Id} were higher than those obtained by Bonenberger at 0 °C. We believe this difference is due largely to the method used to prepare the specimens. Precompression used by Bonenberger probably elevated the yield strength while suppressing the toughness. Also the precompression produced a residual tensile stress near the notch that was only partially relieved by the formation of a short crack at the crack tip. Accordingly Bonenberger measured K_{Id} values (40 to 50 MPa·$m^{1/2}$)

that were significantly lower than the average crack arrest toughness (83.4 MPa·m$^{1/2}$). In this investigation the measurements of K_{Id} at room temperature varied from 58.2 to 95.4 Mpa·m$^{1/2}$ with an average value of 69.6 MPa·m$^{1/2}$, which better agrees with the arrest toughness of the material.

We expect the arrest toughness of the steel K_{Ia} to be less than the dynamic crack-initiation toughness K_{Id}. In this series of measurements, the average value of K_{Id} was about 16.5 % lower than the average value of K_{Ia}. However, in the round-robin testing [7] of the same lot of A533 Type B steel the standard deviation for the crack arrest toughness was 10.6 MPa·m$^{1/2}$. If one defines the lower-bound toughness as the mean less two standard deviations, the value from the round-robin test program would be 83.4 - 2(10.3) = 62.2 MPa·m$^{1/2}$. It appears that the results from the sharply notched modified Charpy specimens are somewhat conservative when average values are considered. However, if one is attempting to determine the lower-bound toughness to be used in a conservative design, then the results using the proposed method appear to be in reasonable correspondence with arrest-toughness measurements when the dispersion of the test results are taken into account.

It is important to eliminate the need to use strain gages in developing a measurement method for fracture toughness that is based on impact loading with Charpy machines. The strain gages require additional instrumentation, are expensive to install and require additional time in data reduction. Master curves can be developed for different alloys using a limited number of strain gages and then employed in subsequent certification testing to insure that the toughness of a particular heat of steel exceeds a minimum lower-bound toughness. This approach requires that the initiation time be established with some degree of confidence. We have shown that the instrumented tup gives a reasonably close estimate of the initiation time. However, before we close this topic it should be noted that initiation is not an instantaneous event.

When a crack initiates, even from a fatigue-sharpened crack, the large-scale initiation event is a series of many small initiation events. Initiation occurs at many separate, small cleavage initiation sites. Crack-front stretch and even limited amounts of tearing (ductile hole-joining) occur before all of the cleavage initiation sites are triggered. Sometimes in the tougher steels cleavage will extend across a few grains and then the fracture mechanism will transition back to ductile hole-joining. The initiation process is a combination of many local initiation events that occur over a small interval, but not at the same instant. When we attempt to measure the initiation time, we should recognize that it is not a precise number, but an approximation of the time when a critical number of local cleavage initiation events are occurring over a finite duration.

ACKNOWLEDGEMENTS

James W. Dally appreciates the support of the National Science Foundation during the later phase of this study.

REFERENCES

[1] Rosenfield, A.R. and Marschall , C.W., "Fracture Mechanics Based Failure Analysis," USNRC Report NUREG/CR-5860, June 1992.

[2] Anon, 1993 Annual Book of ASTM Standards, Section 3, Metals Test Methods and Analytical Procedures, ASTM E-399 and ASTM E-1221 Specification, ASTM, American Society of Testing Materials, Philadelphia, 1993.

[3] Link, R.E. and Joyce, J.A., "Experimental Investigation of Fracture Toughness Scaling Models," 2nd Symposium on Constraint Effects in Fracture, ASTM, November, 1993.

[4] Bonenberger, R.J. and Dally, J.W., "Further Studies of a Modified Charpy Specimen for Lower Bound Toughness," Fracture Mewchanics: 25 th Volume, ASTM STP 1220, F. Erdogan and R.J. Hartranft, Eds., American Society of Testing Materials, Philadelphia, 1993.

[5] Bonenberger, R.J., Dally, J.W., and Irwin, G.R., "Lower Bound Initiation toughness with a Modified Charpy Specimen," Constraint Effects in Fracture, ASTM STP 1171, E.M. Hackett, K.-H. Schwalbe, and R.H. Dodds, Eds., American Society of Testing Materials, Philadelphia, 1993, pp. 139-157.

[6] Winkler, S. Fraunhofer - Institut für Werkstoffmechanik 7800 freiburg, Germany, Private Communication.

[7] Barker, D.B., Chona, R., Fourney, W.L., and Irwin, G.R., A Report on the Round Robin Program to Evaluate the Proposed ASTM Standard Test Method for Determining the Plane Strain Crack Arrest Fracture Toughness, K_{Ia}, of Ferritic Materials, NRC report, NUREG/CR-4996, ORNL/Sub/79-7778/4, January 1987.

[8] Irwin, G.R., Dally, J.W., Zhang, X-J, and Bonenberger, "Lower Bound Initiation Toughness of A 533B Reactor Grade Steel," Symposium on Rapid Load Fracture Testing, San Franciso, CA, April 1990, ASTM, Philadelphia, PA, 1990.

[9] Impact Testing of Metals, ASTM STP 466, ASTM, Phildelphia, PA, 1970.

[10] Radon, J.C. and Turner, C.E., "Fracture Toughness Measurement by Instrumented Impact Test," Engineering Fracture Mechanics, Vol. 1, 1969, pp. 411-428.

[11] Server, W.L. and Tetleman, A.S., "The Use of Pre-cracked Charpy Specimens to Determine Dynamic Fracture Toughness," Engineering Fracture Mechanics, Vol. 4, 1972, pp. 3673-375.

[12] Server, W.L., "Impact Three-Point Testing for Notched and Precracked Specimens," Journal of Testing and Evaluation, Vol. 6, No. 1, 1978, pp. 29-34.

[13] Hoyt, S.L., "Notched-Bar Testing and Impact Testing," Proceedings ASTM Part II, Vol. 38, ASTM, Phildelphia, PA, 1938, pp. 141-156.

A. Karl Schmieder,[1] Patrick T. Purtscher,[2] and Daniel P. Vigliotti[2]

The Role of Strike Marks on the Reproducibility of Charpy Impact Test Results

REFERENCE: Schmieder, A.K., Purtscher, P.T., and Vigliotti, D.P., "The Role of Strike Marks in the Study of Reproducibility of the Results of Charpy Impact Tests," Pendulum Impact Machines: Procedures and Specimens for Verification, ASTM STP 1248, T.A. Siewert and A.K. Schmieder, Eds., American Society for Testing and Materials, Philadelphia, 1995.

ABSTRACT: Charpy V-notch specimens from one lot of high-strength steel were tested by use of three machines to determine reference values for three measures of toughness: absorbed energy, lateral expansion, and height of shear lips. The broken specimens were examined to determine the location and magnitude of changes in specimen features made during testing. The features of interest were the height and location of the shear lips, the location of the lateral expansion projections ranked by height, and the location, length, width, and angle of the first- and second-strike marks. Changes in these features were compared to the changes in average absorbed energy for each of the machines in its standard condition. To correlate changes in these features with intentional machine modifications, ten series of tests were made on a fourth machine. Patterns that could predict the direction of change in absorbed energy for most modifications were observed. The trends indicated by these data are: (1) each modification resulted in an increase in absorbed energy, (2) the distance between second-strike marks is a measure of the compliance of the striking edge and anvil, (3) the offset of the first-strike marks is largely due to lift-off of the specimen at the moment the striker hits the specimen, (4) offset and the angle of second-strike marks are measures of general asymmetry of loading, and (5) lateral expansion and shear lips are valuable as means to detect scale errors and excess losses not related to the work of fracture.

KEYWORDS: Absorbed energy, Charpy V-notch, high-strength steel, high-speed photography, impact machine, lateral expansion, shear lips, strike marks.

[1]Mechanical Testing, Scotia, NY 12302

[2]NIST, Boulder, Colorado 80303

NOMENCLATURE

composite postfracture energy – a measure of the work done to create the fracture surface, excluding energy expended in shock losses, toss losses, and work to form depressions in the specimen at points of contact.

lift-off – the momentary loss of contact between the specimen and the anvils, which occurs immediately after first contact.

offset – the horizontal distance from the striking edge to the notch.

reference value – a value obtained from tests made by a machine in standard condition.

specimen locations – the location of the specimen when in the position for testing.

> In direction of swing:
> (N)[3] notched surface – the surface parallel to the notch root and nearest it.
> (S) struck surface – the surface parallel to the notch root and farthest from it.
>
> vertically:
> (U) upper horizontal surface.
> (L) lower horizontal surface.
>
> transversely:
> (I) inboard half – portion of specimen originally between the striking edge and the machine column.
> (O) outboard half – the portion originally adjacent to, but not including, the inboard half.

first-strike mark – the mark made on the specimen by the anvils before fracture (see Fig. 1).

second-strike mark – the mark made on the specimen if the broken halves fly away from the pendulum and strike the anvils (see Fig. 1).

third-strike mark – additional mark made on the specimen after fracture by striking a machine part or other solid object in the vicinity.

types of second strike marks:

> line – the second-strike mark that completely crosses the notched surface (see Fig. 2, codes A, I and J).
>
> nick – a second-strike mark that impacts both edges, but not the center, of the notched surface (see Fig. 2, codes B and F).

---

[3]Letters are abbreviations used in Tables.

others – a nick at one edge only, or a nick on one half and a line on the other half, or no second-strike marks on one half.

INTRODUCTION

Relative to hardness and tension tests of metal, impact tests have poor reproducibility. This is economially important since it requires more tests to ensure a given degree of precision. This testing problem has been long recognized [1-4]. Although significant improvements have been made during the last three decades [5], limited reproducibility remains one of the principal disadvantages of impact testing. The objective of this study is to reduce the variability of impact test results by identifying machine deficiencies through inspection of the broken specimens.

This inspection included the usual measurement of absorbed energy, lateral expansion, and shear lips. In addition, another less well-known measurement was made, the characterization of subsequent strike marks. The process of forming the marks is shown schematically in **Fig. 1**. High-speed photographs confirm the transverse flight of the brokenhalves and the strike against the anvils. Enlarged photographs in **Fig. 2** show various types of first- and second-strike marks.

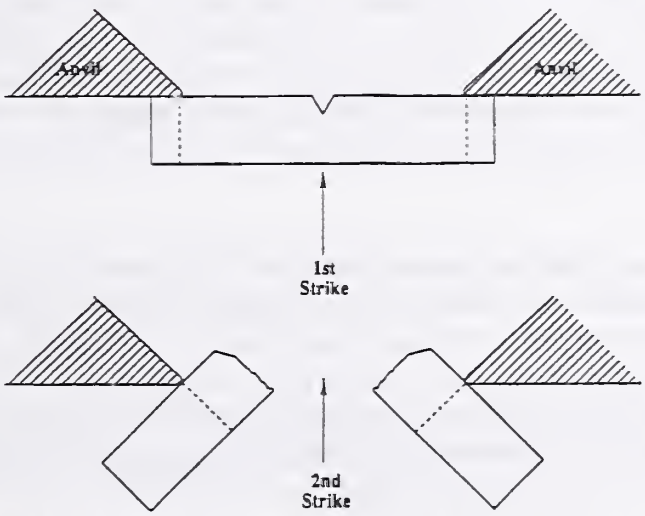


Figure 1. Schematic diagram shows how the first- and second-strike marks are produced.

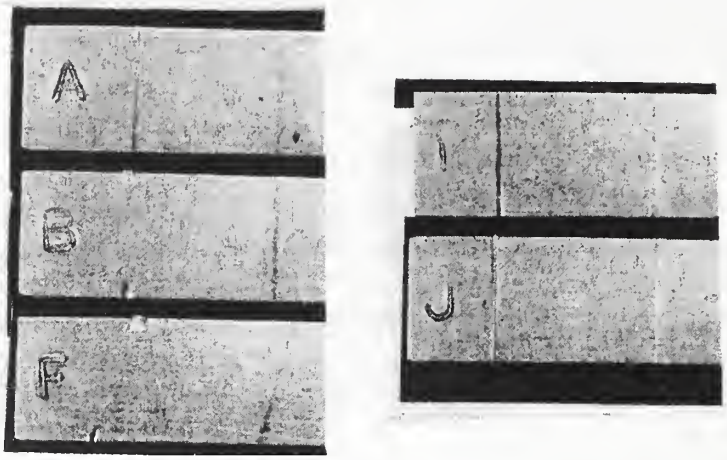


Figure 2. Photographs of the struck surfaces on a broken half from each of the following series: A, B, F, I, and J. The first-strike marks are on the right-hand side and the second-strike marks are on the left-hand side.

EXPERIMENTAL PROCEDURES

Specimens

All specimens were drawn at random from a large lot of verification-grade specimens. The material specified is published in ASTM Standard Practice for Qualifying Charpy Verification Specimens of Heat Treated Steel, E 1271, Appendix X1. The specimen dimensions were those for Type A shown In Fig. 6 of ASTM Standard Methods for Notched Bar Impact Testing of Metallic Materials E 23, except that the tolerances are smaller.

Machines

The first three machines, that were used to determine the reference values were manufactured by different companies to meet the specifications of ASTM E 23. All had capacities of 300 J (220 ft·lbf) or more. All were directly verified within a year of making the tests reported here. The fourth machine, which was modified during these tests, had less than half the capacity of the others. The machine designations used in this study, the pendulum types (as described in ASTM E 23, Fig. 1), and the scale errors in percentage of the reading are:

Designation	R1C	R1U	R2U	M1C
Type of Pendulum	C	U	U	C
Scale error	-3.0%	-0.8%	-0.3%	0.0%

All reported values are corrected for the scale errors that are known. A method to estimate the scale error for M1C is discussed later. The modifications made to Machine M1C are given letter designations and are listed below. The same letters are used to identify the test series in the tables of results.

A. As received. History of prior use unknown. Tightness of bolts unknown.
B. Old anvils replaced by new anvils. Bolts tightened to 100 J (75 ft·lbf) at each installation.
C. Old anvils reinstalled.
D. Place specimen on the supports so that they are offset 2 mm toward the machine pedestal.
E. New taller supports installed to raise the specimen 10.6 mm above the standard position. During this and each subsequent replacement of the supports, the bolts were tightened to 27 J (20 ft·lbf).
F. Reinstall the original inboard support only.
G. Remove both supports so the specimen is 10.9 mm below the standard position.
H. Shorten and grind the top surface of the new supports so that the specimen is slanted upward toward the anvil at one degree with the horizontal; reinstall supports.
I. Reinstall the original standard supports. Remove old anvils and grind the face that bears against the specimen so that it has an angle of 10:1000 to the original surface, measured in a vertical plane when installed. Restore the corner radii and surface finish. Reinstall modified anvils.
J. Remove anvils and restore contact faces to original condition; also reduce the thickness in the direction of swing by 5 mm. Reinstall modified anvils.
K. Reinstall new anvils, restoring the machine to standard condition. Photograph specimen half as it flies transversely and strikes the anvil.

Methods of Testing

All specimens were tested at −40°C (−40°F) in accordance with ASTM E 23. The lateral expansion of the broken specimens was measured as prescribed in ASTM E 23-93a, Section 12.4.2.

The location, length, and width of the strike marks were measured with an optical comparator in the reflective mode. Magnifications of 10X and 20X were used for the first- and second-strike marks, respectively.

Many of the first-strike marks had poorly defined outer boundaries. To be consistent, we reported the distance from the notch root center-line to the inner boundary of the mark. For a few marks, because both boundaries were poorly defined no dimension was recorded.

The high-speed photographs were taken at 2000 frames per second with a video system. The lighting system included two banks of incandescent lights arranged around the outboard side of the impact machine and drawing a total of 1500 watts of power.

Methods of Calculation

The statistical calculations were made according to the mathematical definitions of the average and standard deviation. The lateral expansion was measured as prescribed by ASTM E 23, that is, by the sum of the two highest of the projections at the ends of the pendulum strike marks on both broken halves. This method is based on these results that correlate lateral expansion and absorbed energy over a wide range, mostly at energies higher than those reported here. Some information is available in the range of these tests [6].

Since four measurements are taken in any case, they were added to see whether the correlation with absorbed energy was improved in the range of these tests. To make the results compatible with the standard value, the sum of the four measurements was divided by two, then reported.

The composite postfracture energy index reported was calculated by dividing the average value for a series by the average of all specimens tested during the program. This was repeated for each of the three types of measurements of postfracture energy, and then the three values were averaged. The result is a dimensionless number for each series that allows the series to be compared but does not indicate the magnitude or units of the energy measurement. To allow comparison to other published results, the composite number was multiplied by the average value for all tests for the lateral expansion measured according to ASTM E 23. The average values for all tests are shown on the bottom line of **Table 1** and the values for Series R1C on the top line. Using these values, the composite postfracture index for Series R1C is

$$0.133\left[(0.12/0.133)+(0.11/0.111)+(1.35/1.403)\right]/3 = 0.126\,mm$$

This distance, reported as a measure of position of second-strike marks, is the distance from the centerline of the notch to the centerline of the mark. The offset is equal to one half of the difference between first-strike marks, measured on the inboard and the outboard halves. Unless stated otherwise, the distances at the upper and lower surfaces are averaged before the difference is calculated.

RESULTS

Table 1 shows averages and standard deviations of energy measurements for some single series and for combinations of related series. **Table 2** shows the energy measurements for all of the series as a dimensionless ratio to the average values for Machine M1C. This permits comparison of quantities such as lateral expansion to absorbed energy that have different units and magnitudes varying by a factor of over one hundred.

Table 3 presents the various second-strike measurements. **Table 4** is a tally showing the number of occurrences of various deformations at specified locations. Its primary use is to identify asymmetrical conditions. **Table 5** presents statistics on first- and third-strike marks and compares absorbed energy to the composite postfracture energy.

Table 1. Weighted average values for quantities which are given as a single number for both broken ends

Machine or Class	Absorbed Energy J, ft·lbf	Lateral Expansion[a]		Sum of Shear Lips, mm
		E 23, mm	Sum/2, mm	
R1C	15.9, 11.7 (0.3%)[b]	0.12 (0.01)[c]	0.11 (0.01)[c]	1.35 (0.11)[c]
R1U	17.2, 12.7 (0.5)	0.14 (0.03)	0.12 (0.0●)	1.43 (0.14)
R2U	17.5, 12.9 (0.5)	0.14 (0.03)	0.12 (0.01)	1.43 (0.14)
M1C-A	15.2, 11.2 (0.6)	0.11 (0.10)	0.●● (0.01)	1.34 (0.04)
A,B,C,K	15.3, 11.3 (0.2)	0.13 (0.02)	0.10 (0.01)	1.34 (0.04)
M1C-Others	15.9, 11.7 (0.3)	0.14 (0.01)	0.11 (0.01)	1.45 (0.09)
All RXX	16.9, 12.4 (0.6)	0.13 (0.01)	0.12 (0.●●)	1.39 (0.04)
All XXU	17.4, 12.8 (0.2)	0.14 (0.00)	0.12 (0.00)	1.40 (0.04)
All Std.C[d]	15.5, 11.4 (0.2)	0.13 (0.00)	0.10 (0.00)	1.34 (0.01)

[a] E 23 is sum of two highest projections. "Sum/2" is one half the sum of the four readings.
[b] Values in parenthesis are coefficients of variation.
[c] Values are standard deviation in mm.
[d] Values from R1C and M1C: Series A, B, C, K.

Table 2. Deviations[a] of energy-related measurements in Table 1 from the corresponding values for Machine R1C.

ID[b]	Absorbed Energy	Lateral Expansion		Height of Shear Lip
		E 23 method	Sum of Four	
R1C (10)	2.6 (0.6)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
R1U (10)	8.5 (0.7)	14.6 (2.7)	12.5 (0.9)	15.9 (0.3)
R2U (10)	10.3 (0.7)	15.2 (1.6)	15.9 (0.6)	2.4 (0.3)
M1C-A (6)	-4.5 (1.1)	-6.3 (0.7)	-14.0 (0.3)	0.7 (-0.2)
B (6)[c]	-●.● (0.6)	5.7 (5.4)	-1.5 (3.5)	0.7 (0.5)
C (5)[c]	-3.0 (0.2)	13.3 (-0.6)	15.1 (-0.4)	-4.8 (0.1)
● (5)	2.6 (0.6)	28.5 (0.2)	1.1 (0.4)	14.1 (0.7)
E (5)	-1.7 (0.5)	5.3 (1.4)	-4.2 (0.7)	3.0 (-0.3)
F (6)	-2.7 (0.6)	7.8 (-0.4)	-4.5 (0.5)	3.0 (0.1)
G (5)	-0.9 (0.6)	12.1 (2.8)	6.8 (1.3)	3.0 (0.9)
H (5)	1.7 (0.0)	18.4 (1.4)	13.9 (0.2)	4.4 (0.9)
I (5)	1.3 (1.1)	20.5 (0.6)	4.6 (0.8)	0.7 (-0.2)
J (5)	3.4 (0.2)	18.4 (3.3)	16.3 (0.2)	10.8 (0.2)

[a] Deviation of measured quantities are shown as percentages, deviation of the coefficient of variation of that quantity as a ratio.
[b] Machine identification, series, and number of specimens.
[c] One of the halves not available for measurement.

Table 3a. For the reference machines, average values and coefficients of variation for positions, width, and direction of second-strike marks.

ID[a]	Distance from Notch to Mark[b]				Line Width,[c] mm	Angle with Upper Edge, deg.			
	Outboard		Ratio 0/I			Outboard		Inboard	
	Lines	Others	Lines	Other		Lines	Others	Lines	Others
R1C	0.94 (0.04)[d]	0.94 (0.04)	1.03 (0.08)	1.06 (0.09)	1.22/1.12[e] (0.38)	0.3 (0.7)	3.2 (0.4)	0.2 (1.0)	0.5 (0.4)
R1U	0.94 (....)	0.93 (0.04)	0.98 (....)	1.03 (0.07)	0.28/0.34 (0.31)	1.0 (...)	2.0 (0.3)	0.2 (...)	0.2 (0.5)
R2U	1.06 (0.02)	0.99 (0.04)	1.05 (0.06)	0.99 (0.03)	1.25/1.22 (0.37)	1.6 (0.5)	2.7 (0.8)	0.8 (1.3)	0.2 (1.1)

[a] Machine identification and type of pendulum.
[b] Dimensionless ratio of distance to width of specimen, which is 10 mm.
[c] Width of all specimens in a series combined statistically.
[d] Value in parenthesis is ratio of standard deviation to average.
[e] Upper surface/lower surface.

DISCUSSION

Significance and Limitations of Various Measurements

One of the main objectives of the impact test was to measure the energy required to produce the fracture surface. The loss of potential energy of the pendulum during the swing is reported as absorbed energy, but it also includes:

A. frictional losses due to pendulum motion,
B. shock losses due to vibration and displacement of the machine parts,
C. crushing work to form the depressions on the surfaces struck by the pendulum and the anvils simultaneously, and
D. the kinetic energy of the broken halves after fracture.

Table 3b. For Machine M1C, average values and coefficients of variation for position, width, and direction of second-strike marks.

ID[a]	Distance from Notch to Mark[b]		Line Width,[c] mm	Angle with Upper Edge,	
	Outboard	Ratio O/I		Outboard	Inboard
A	0.94 (0.02)	1.01 (0.05)	0.49 (0.16)	1.3 (0.8)	1.2 (1.1)
B	0.90 (0.06)	1.01 (0.05)	0.65 (0.30)	2.5 (0.4)	2.1 (0.7)
E	0.90 (0.07)	1.00 (0.07)	0.57 (0.50)	1.5 (0.9)	1.9 (1.0)
D	1.09[e] (0.01)	... (...)	0.56[e] (0.11)	1.6[e] (0.4)	... (...)
E	0.89 (...)	0.95 (...)	0.64 (0.10)	1.5 (0.9)	1.4 (0.8)
F	0.83[f] (0.05)	... (...)	1.21[f] (0.10)	... (...)	19.0 (0.3)
G	0.86 (0.04)	0.95 (...)	0.67 (0.29)	2.6 (0.4)	... (...)
H	0.86 (0.05)	0.99 (0.05)	0.84 (0.28)	2.8 (0.7)	4.7 (0.4)
I	0.89 (0.03)	1.05 (0.06)	0.56 (0.15)	1.3 (0.9)	1.1 (0.6)
J	0.86 (0.06)	1.04 (0.11)	0.49 (0.17)	1.8 (0.8)	1.0 (0.7)

[a] Series identification.
[b] Dimensionless ratio of distance to width of specimen, which is 10 mm.
[c] Width of all specimens in a series combined statistically.
[d] Value in parentheses is ration of standard deviation to average.
[e] Outboard values. No strike marks on inboard halves.
[f] No strike marks on upper outboard halves.

Table 4. Evaluation of machine asymmetry by number of deformations at various locations.

ID[a]	Start of Second-Strike Marks[b]						Shear Lips		Lateral Expansion Projections[c]							
	Line		Nick		None[d]				Upper Surface				Lower Surface			
	U	L	U	L	U	L	U	L	A	B	C	D	A	B	C	D
R1C																
-O	6/1[e]	6/4	4/2	4/3	0	0	5/2	4/3	2	3	1	4	3	2	2	3
-I	7/2	7/5	3/1	2/1	0	1	5/3	6/2	2	4	1	3	4	4	1	1
R1U																
-O	6/6	6/0	2/0	4/4	2	0	3/2	6/1	2	2	1	3	4	2	1	1
-I	3/1	3/2	6/2	6/5	1	1	7/4	4/3	1	3	2	2	1	3	3	1
R2U																
-O	6/6	6/0	4/1	4/3	0	0	5/3	5/2	5	2	2	1	3	3	4	0
-I	6/5	6/1	4/3	4/1	0	0	5/3	5/2	3	3	2	2	1	2	0	7
M1C																
A-O	2/2	2/1	2/1	3/1	1	0	4/2	3/1	1	2	2	1	3	3	0	0
A-I	6/4	6/2	0	0	0	0	3/2	3/0	2	1	1	2	2	0	1	3
M1C																
B-O	2/2	2/0	2/1	3/1	1	0	2/1	0	2	1	2	0	2	0	1	2
B-I	1/1	1/0	4/0	5/0	1	0	3/1	6/3	1	1	1	2	2	1	1	1
M1C																
D-O	5/5	5/0	0	0	0	0	1/0	4/1	0	1	1	2	0	3	1	0
D-I	0	0	0	0	0	0	4/3	1/1	0	0	1	3	1	2	0	1
M1C																
H-O	3/2	3/1	2/1	2/1	0	0	3/2	3/2	2	1	2	0	0	3	2	0
H-I	1/0	1/1	3/1	4/2	1	0	2/0	2/1	2	0	1	2	1	1	2	1
M1C																
J-O	3/3	3/0	2/2	2/0	0	0	2/1	3/2	1	1	0	3	3	1	1	0
J-I	5/3	5/2	0	0	0	0	3/1	2/1	1	0	2	2	1	2	2	0

[a] Machine identification - Position of half before fracture.

[b] Where first contact between the specimen and the anvil occurs, "U" designates uppermost surface in positioning for testing, "L" the opposite.

[c] Letters indicate ranking of heights at four locations. A is highest projection. If two are equal, both receive the same rank and the next lowest is omitted.

[d] Indicates no second strike mark at that location.

[e] Numerator is total number of occurrences at that location. Denominator is the number of highest values included in the numerator.

Table 5. Absorbed energy, index of fracture work, first-strike dimensions, and frequency of occurrence of third-strike marks.

ID	Absorbed Energy, J	Composite Index of Fracture, mm	First-Strike Marks			Third-Strike Marks[a]			
			Offset	Angle, deg		Fracture End			
			mm	O	I	Lips[b]	Corner	Surface	Edge
R1C	15.9 (2.5)[c]	0.126	0.1 (0.5)	0.4° (0.7)	0.2° (1.5)	67	10	7	27
R1U	17.2 (4.0)	0.140	0.2 (0.7)	0.4 (1.5)	0.1 (2.1)	55	22	11	27
R2U	17.5 (3.9)	0.140	0.0 (0.4)	0.2 (1.8)	0.4 (1.0)	50	...	...	45
- A	15.2 (5.6)	0.113	0.1 (0.9)	0.6 (0.9)	0.2 (1.6)	109	20	55	45
- B	15.0 (3.6)	0.128	0.1 (0.8)	0.1 (...)	0.9 (1.6)	110	20	80	10
- C	15.4 (3.0)	0.140	0.1 (0.4)	0.3 (1.2)	0.9 (1.6)	80	22	67	44
- D	16.3 (3.9)	0.140	2.1 (1.3)	0.2 (1.4)	0.1 (2.2)	0	0	0	0
- E	15.6 (3.7)	0.128	0.0 (0.4)	0.2 (1.4)	0.2 (1.4)	55	10	36	64
- F	15.4 (3.0)	0.127	0.5 (0.5)	12.8 (0.1)	14.3 (0.1)	40	20	40	10
- G	15.7 (4.0)	0.135	0.1 (0.4)	1.1 (0.7)	1.3 (0.8)	83	8	50	33
- H	16.1 (2.5)	0.140	2.1 (1.3)	0.3 (0.9)	0.5 (1.1)	70	...	70	20
- I	16.1 (5.2)	0.140	0.1 (0.9)	0.3 (0.9)	0.6 (0.7)	67	...	44	44
- J	16.4 (2.8)	0.145	0.2 (0.5)	0.2 (1.4)	0.3 (0.9)	22	22	22	22

[a] Values shown are number of occurrences as a percent of the number of specimen halves examined.

[b] Flattening of the tip of the shear lip, thereby reducing the height.

[c] Values shown in parentheses are standard deviations as percentages of the value.

Absorbed energy – ASTM E 23 includes a correction for (a) that adequately removes this amount of work from the reported absorbed energy [7]. The other losses are included in the reported value. Measurements of lateral expansion remove all of the above except the crushing work at the struck surface. The shear-lip-height method excludes all of the losses. However, that does not mean that the machine has no influence on the results from these alternative measures of energy. These tests indicate that the different machines impose different conditions of force, displacement, and loading points that influence the work done at or near the fracture; therefore, the resulting numbers are not solely material properties. Attempts to separate these machine-dependent contributions by measurements of the broken specimen halves is discussed further in other sections.

Lateral expansion – The principal use of lateral expansion in this study is to provide a measure of the work to produce a fracture without depending on the energy measurement scale of an impact machine. In order to estimate its discriminatory capability in a single series, a regression analysis was made with inputs of the average absorbed energy and lateral expansion from each of the fourteen series unweighted for the number of specimens in each series. The regression of series averages showed regression coefficients 0.534 for the ASTM method, 0.644 for the sum method, and 0.552 for the composite postfracture energy.

Shear-lip heights – The shear-lip height has the advantage of being subject to less proportional reading error than lateral expansion; the shear-lip heights were about 10 times greater than the lateral expansions. This difference would seem to compensate for the differences in the reference surfaces for the two methods, a fracture surface for the former and a ground surface for the latter. Another disadvantage of the shear-lip height is that a maximum is sought along the crest of the lip by measuring at various distances from the notch. A much more serious source of error is that most crests were flattened by various amounts due to contact with a hard plane surface.

A regression analysis of the ten individual specimens vs. the absorbed energy of each showed that the shear-lip height was superior to the lateral expansion method for a single series. The regression coefficients were 0.81 for the R1C machine, 0.56 for R1U, and 0.46 for the R2U machine. Corresponding values for lateral expansion were all less than 0.25.

First-strike marks – The accuracy of the measurements can be judged by comparing the distance between first-strike marks on the inboard and outboard halves to the specified distance between anvils. For the three reference machines, the distance between marks are 41.0 ± 0.2 mm for R1C, 41.1 ± 0.1 mm for R1U, and 41.1 ± 0.2 mm for R2U. The specified gap between anvils is 40.0 ± 0.05 mm and the distance between tangent points (gap + radius of curvature) is approximately 42.0 mm The average of the values for the three machines differs from the average of the specified dimensions by 0.2 mm of 0.5 percent, indicating that the distance measurements have better precision than any other measurements presented in this report. The angle measurements are derived trigonometrically from the distance measurements. Their accuracy is estimated at 0.3 deg.

For the thirteen series of tests reported in **Table 5**, the average offset of first-strike marks was typically less than 0.3 mm; four values are less than 0.1 mm, indicating that the specimens had been positioned very precisely. One explanation for higher offsets is operator carelessness, but that would not account for the fact that the angles, which the operator does not control, vary just as much. Even more puzzling is the fact that the angles on each end of the specimen differ, in some cases, by more than the estimated inaccuracy. An explanation is that after the first contact with the striker, the specimen bends and loses contact with the anvils momentarily. Furthermore, when contact is restored, one anvil may touch before the other, allowing the specimen to continue tilting until the second anvil makes contact. The published records of force vs. time during the first contact between the specimen and the striking edge of a high-strength steel specimen show the force decreasing, approaching zero for a short period, thus indicating that lift-off occurs [8,9].

Second-strike marks – The second-strike marks provide more measurements related to asymmetry that the other reported characteristics of the broken halves. The most conspicuous is whether the marks are lines or nicks. A less obvious but just as useful difference is the width of the strike mark, the edge showing the widest mark having struck the anvil first.

The feature found most useful for quantitative analysis is the sum of the distance between second-strike marks, obtained by adding the distance from notch to mark for the outboard half of the specimen to that for the inboard half. This distance seems to be a measure of compliance of anvils and frame. The test series did not contain a comparison of directly controlled compliance, but Series I an J have partial contact between the specimen and the anvils during the initial loading, thus simulating the slower rate of increase in force of a less stiff system. The average distance between marks for the same machine and the same anvils in normal condition (Series A and C) was 36.9 mm, while that for the less stiff anvils (Series I and J) was 35.8 mm, or a difference of 3 percent. The corresponding increase in absorbed energy was 15.3 to 16.2 J or an increase of 3 percent, both changes being larger than those due to other modifications. Bluhm has demonstrated that the absorbed energy increases with a decrease in machine stiffness [10].

Third-strike marks – For Type-C machines, the specimen halves can leave the machines transversely without again being near the moving pendulum. For Type-U machines, the specimens rebound from the shrouds and approach the moving pendulum. With this in mind, we expected that the Type-U machines would have more third-strike marks. This expectation is not confirmed by the test results shown in **Table 5**.

The reported effectiveness of shrouds in reducing jamming was confirmed by these tests [5]. Only one specimen half showed marks due to jamming. Even in this case, an increase in absorbed energy was not clearly demonstrated. One specimen of the ten in the series had greater absorbed energy than the jammed specimen.

Evaluating asymmetry of machines – At the beginning of this study, we assumed that a properly adjusted machine would produce a uniform line as a second-strike mark on the specimen. Misalignment would result in asymmetrical second-strike marks. The data indicate that the absorbed energy increases as deliberate asymmetry is introduced, but the rate of increase for a given asymmetry is small. Furthermore, the marks and features of the broken halves are also asymmetrical. When the specimens show second-strike marks, there are several features that can be used to measure symmetry. Among these are:

A. whether lines or nicks are formed on the two halves,
B. the distance from the notch on each half,
C. the angle of the strike mark with the specimen edge, and
D. whether the upper or lower edge struck the anvils first.

Table 4 shows the tally for these dimensions. Examples of how the second-strike marks are useful will be presented in the discussion of the results that follow.

Components of absorbed energy – The simplest estimate of the effect of a modification is to subtract the absorbed energy of tests with the modification from that without the modification. This is effective if other variables not obviously related to the controlled variable remain unchanged. To decrease the proportional effect of the uncontrolled variables, the controlled variable was changed by an amount five to ten times the specified tolerance on the controlled variable. An exception is for specimen elevation relative to the striking edge, which could be changed only to the limits of the specified tolerance. Then the difference in absorbed energy was divided by the change in the controlled variable to measure the rate of change, assuming a linear relationship. For example, the effect of offset was calculated to be +0.44 J/mm, that of the angle of strike to be 0.02 J/deg, and that of compliance as measured by the change in the distance between second-strike marks to be −0.92 J/mm.

The second approach was to make a regression analysis of individual specimens from a set which showed an above average range of that variable. That approach showed much larger values of slope; for example, 4.0 J/mm for offset. The 4.0 J/mm estimated the change between some modifications better than 0.44 J/mm. However, the higher value is not due to offset alone; other uncontrolled factors must be present that influence the result that we measure.

Comparison of Pendulum Types

Table 2 shows that the average value for absorbed energy measured by R1U and R2U is 9 percent greater than that measured by R1C. This is greater than the change in absorbed energy due to deliberate modifications made to exceed the specified tolerance beyond what would be expected in machines in use. Considering the magnitude of difference due to pendulum type, this effect will be analyzed in detail.

Comparison of different measures of the energy to produce fracture – As shown in **Table 5**, on average the composite postfracture energy, after testing is 10 percent higher for tests made using the Type-U machines than for those using the Type-C machine. Each method of measurement in Table 1 indicates that the Type C machines produce lower values than for the Type-U machines. Therefore, the difference in absorbed energy is not due to the means of measurement.

Differences in the striking edge – We know that the stiffness of the loading system affects the absorbed energy and that the difference can be measured fairly easily [8,10]. The striking edge of the type-C pendulum is supported along its whole length by the massive disk of the beam. The stiffness of the latter is no doubt less than the former so the rate of increase of the force on the specimen is smaller and less uniform from the upper side of the specimen to the lower.

Another factor to consider is the distance between second-strike marks. For a specimen that is loaded more rapidly, we expect that the halves would have greater velocity after fracture, and thus the second-strike marks would be closer together. The distance between second-strike marks for the specimens broken in the Type-C machine average 19.0 mm; those broken by R1U average 18.6 mm; and those from R2U average 21.8 mm. The average difference of 2.8 mm between machines R1C and R2U would lead to a predicted increase in energy of about 16 percent for R2U compared to R1C; a 10 percent increase was measured (see **Table 2**). The results for R1U do not follow the generally observed trend.

Energy loss due to shrouds – The more complicated flight path for specimens tested in shrouded machines would be a convenient explanation for an increase in the number of third-strike marks when compared to the Type-C machine. However, the tallies in Table 5 show that the numbers are approximately equal. In any case, unless the marked specimen actually makes contact with the pendulum, it cannot affect the absorbed energy.

Summary – Of the three uncontrolled variables (energy loss to shroud, difference in striking edge, and means of measurement), the difference of striking edge appears to be the major factor that relates to the difference between pendulum type.

Effects of Modifications to Machine M1C

The scale of the modified machine was graduated in ft·lbf, and the readings were recorded to the nearest 0.25 ft·lbf, which is about 2 percent of the typical value obtained during the tests described in this paper. Modifications producing changes of greater than 2 percent are regarded as significant and are discussed below.

Offsetting specimen 2 mm – Series D shows that the effect of the offset on the different measures of toughness (see **Tables 2 and 5**) was measurable. In **Table 4**, the effect of the offset on appearance of the broken halves is consistent. None of the inboard halves (shorter from anvil to notch) showed second-strike marks. All of the outboard halves had second-strike lines. The 2 mm offset shows the most conspicuous relationship between the type of modification and the appearance of the broken specimens.

Relative to the normal condition of the machine is Series C, the effect of the offset is to increase the absorbed energy by 0.9 J or about 6 percent, an amount exceeded by only one other modification in the series of eleven. The rate of 0.44 J/mm is consistent with previous experience. Again, the rate of 4.0 J/mm quoted earlier in the discussion is not due to the offset, but is rather the result of uncontrolled variables.

Tilting the specimen supporting surface – Series H consists of tests made after the new supports were machined to position the specimen at the standard height but with the horizontal surface sloped upward toward the anvils. The effect on absorbed energy was an increase of 0.7 J or 5 when compared to the machine in its last standard condition. The composite postfracture energy (Table 5) increases slightly from 5.3 to 5.6 mm, indicating that the change is relatively insignificant. Comparing the distance between second-strike marks shows values of 18.1 and 17.5 mm, respectively, for the reference Series C and Series H. Multiplying the difference by −0.92 J/mm (the previously determined slope) predicts a +0.6 J change in absorbed energy due to the reduced stiffness of the modified system. Therefore, the change in absorbed energy can be predicted from the appearance of second-strike marks in this case.

Remachining the anvils to slant the contact faces – Table 5 shows that Series I averages 0.7 J or 4 percent more than Series C. The misalignment of 10 parts in 1000 is about 0.6 deg, less than that tested in Series H, where the support was angled 1 deg. The distance between second-strike marks is 17.8 mm. The 0.3 mm decrease in distance between second-strike marks would account for a 0.3 J increase in absorbed energy, again consistent with the analysis in the previous section.

Reducing the horizontal thickness of the anvils – Table 2 shows that when the anvils are moved in the direction of swing by 5 mm (Series J), the absorbed energy is increased by 6 percent relative to the machine in standard condition (Series C). This is larger than the corresponding value for any other modification made during these tests. The physical effect of this on fracture work is that it causes a change in the angle between the specimen and the striking edge, similar to Series H and I. In this case, the angle is about 5 parts in 1000 and the distance between second-strike marks in 17.5 mm (0.6 mm decrease from that found in series C). The predicted change in absorbed energy is inconsistent with the predictions in the previous two sections. However, the change in absorbed energy is consistent with the change in composite postfracture energy reported in **Table 5**.

CONCLUSIONS

Among the effects of the thirteen variations tested, the largest was a difference between the absorbed energy as measured on a machine with a Type C pendulum and two machines with Type U pendulums, the latter indicating values 9 percent higher. The intentional modifications to M1C indicate that further tightening of machine tolerance would have marginal utility for normal testing. The alternative energy-related measurements were helpful as additional information to substantiate our conclusions. These conclusions may be applicable only to tests on high-strength steels with fracture energies below 20 J.

For similar investigations in the future, we recommend that a set size of at least ten specimens be used, and that the energy indicator on the machine should be readable to 1 percent variations in the absorbed energy.

ACKNOWLEDGMENTS

The authors are grateful to N.V. Gjaja for permission to use the measuring instruments at the Schenectady Materials and Processes Laboratory and to T.P. Quinn and R.B. Madigan (NIST) for assisting with the high-speed photography.

REFERENCES

1. Kahn, N.A., and Imbembo, E. A., "Reproducibility of the Single-Blow Charpy Notched-Bar Test," ASTM Bulletin, May, 1947, pp. 66-74.
2. Fry, N.L., "An Inquiry into the Reproducibility of Impact Test Results," ASTM Bulletin, January, 1953, pp. 61-66.
3. Driscoll, D.E., "Reproducibility of Charpy Impact Test," Symposium on Impact Testing, STP 176, American Society for Testing Materials, Philadelphia, 1955.
4. Rickett, R. L., et al., "A Comparison of Impact Testing Machines in the 20 to 30 Ft-Lb Range," Proceedings ASTM, Vol. 53, pp. 739-754, 1953.
5. Fahey, N.H., "Effect of Variables in Charpy Impact Testing," Materials Research & Standards, November, 1961, pp. 872-876.
6. Orner, G.M., "Charpy Brittle Fracture Transition by the Lateral Expansion-Energy Relationship," Welding Journal, May, 1958, pp. 201s-205s.
7. Schmieder, A. K., "Comparison of Metrological Techniques for Charpy Impact Machine Verification," Charpy Impact Test Factors and Variables, ASTM STP 1072, J. M. Holt, Ed., American Society for Testing and Materials, Philadelphia, 1990, pp. 20-34.
8. Porro, F., Trippodo, R., Bertozzi, R., Garagnani, G., "Impact Tester Compliance: Significance, Sensitivity and Evaluation," Charpy Impact Test: Factors and Variables, ASTM STP 1072, J. M. Holt, Ed., American Society for Testing and Materials, Philadelphia, 1990, pp. 7-19.
9. M.P. Manahan, R.B. Stonesifer, Y. Soong, and J.M. Burger, "Miniaturized Notch Test Specimen and Test Machine Design," Pendulum Impact Machines, ASTM STP 1248, T.A. Siewert and A.K. Schmieder, Eds., American Society for Testing and Materials, Philadelphia, 1995, pp. 39-69.
10. Bluhn, J.I., "The Influence of Pendulum Flexibilities on Impact Energy Measurements," Symposium on Impact Testing, ASTM STP 176, American Society for Testing and Materials, Philadelphia, 1955.

Earl A. Ruth,[1] Daniel P. Vigliotti,[2] Thomas A. Siewert[2]

Effect of Surface Finish of Charpy Anvils and Striking Bits on Absorbed Energy

Reference: Ruth, E. A., Vigliotti, D. P., and Siewert, T. A., **"Effect of Surface Finish of Charpy Anvils and Striking Bits on Absorbed Energy"**, Pendulum Impact Machines: Procedures and Specimens for Verification. ASTM STP 1248. Thomas A. Siewert and A. Kari Schmieder, Eds., American Society for Testing and Materials, Philadelphia, 1995.

Abstract: Some new Charpy impact testing machines often report higher energy values than machines which have been in service for a period of time. An investigation into the cause of this phenomenon revealed that a major contributing factor was the surface finish of the Charpy anvils and the striking bit. Depending on the trajectory of the Charpy specimen halves after impact and the configuration of the testing machine, a considerable amount of friction may occur between the specimen and the anvils as the specimen is broken. Friction may also occur between the specimen and the striking bit as the specimen exits the machine. Polishing the anvils and the striking bit minimizes the friction between these parts. As a Charpy impact testing machine tests many samples, the striking bit and anvils become burnished and energy absorption due to friction is reduced. If new parts are highly polished prior to installation, the burnishing, normally caused by testing many samples, is simulated. In this way energy absorption associated with friction during the wear-in period is minimized and remains more constant during the life of the anvils and striking bit. ASTM Standard Test Methods for Notched Bar Impact Testing of Metallic Materials (E 23) requires the surface finish of the anvil and striking bit to be better than 4 μm (125 μin.). Specifying a 0.1 μm (4 μin.) surface finish as a requirement in E 23 will not only eliminate the wear-in period, but will minimize shifts in test results when anvils and/or striking bits are replaced in Charpy impact testing machines.

Keywords: anvils, Charpy impact test, energy loss, friction, impact test, striker, surface finish

[1] Tinius Olsen Testing Machine Co., Inc., Willow Grove, Pennsylvania 19090
[2] Materials Reliability Division, National Institute of Standards and Technology, Boulder, Colorado 80303

Background

In the fall of 1991, several new Charpy impact testing machines were verified using National Institute of Standards and Technology (NIST) Low-Energy Samples for Charpy V-Notch Machines (SRM 2092). The test results indicated absorbed energies which exceeded the average by more than 1.4 J. An extensive investigation into the cause of the discrepancy was initiated.

All dimensional aspects of the machines were evaluated and were found to be well within the ranges permitted by ASTM E 23. Measurements were made by three different investigators, all of whom developed similar results.

The structural integrities of the machines' frames and foundations were investigated. Experiments were conducted that stiffened the frame. Other experiments that stiffened the pendulum were conducted. No improvements due to these procedures were noted.

One observation made at the time was that the NIST reference specimens (SRM 2092) were tending to exit the testing machine in the direction of swing of the pendulum. In the past, it was typical for these specimens to exit the testing machine in a direction opposite to the direction of swing of the pendulum. Further, the striking bits and anvils on the NIST reference Charpy impact testing machines used to develop the reference value for the SRM's, were burnished to a high polish as a result of having been used to test many specimens. We hypothesized that perhaps the relatively rough surfaces on the new striking bits and anvils were causing additional friction with the specimens as they were being broken. The friction would add to the energy absorbed during a test.

In order to test this hypothesis, new anvils and striking bits were polished to a mirror-like surface and installed on a machine which had previously been unable to pass the NIST requirements on low-energy specimens. A dramatic decrease in energy absorbed was observed at that time. The average energy obtained on the NIST specimens tested on that machine dropped approximately 0.7 J. The tests conducted at that time were not well structured or documented, due to the expediency and shipping schedule demands for the new impact testing machines. As a result, the data from those tests are not included in this paper. A more structured experiment was designed and the results are offered here.

Experimental Method

Charpy V-notch tests were performed in accordance with ASTM Standard Test Methods for Notched Bar Impact Testing of Metallic Materials (E 23). The test temperature in all cases was -40°C. The tests were performed on a 400 J capacity, pendulum Charpy impact testing machine having a "U" type pendulum. The standard size Type A specimen was used. Some of the tests were performed with a 2 mm radius striking bit instead of the standard 8 mm radius striking bit. This machine had been evaluated for repeatability when the strikers were interchanged. The repeatability was found to be better than the uncertainty in the test data.[1] These tests conformed to E 23 in all other ways.

Three types of specimens were tested. The three types of specimens were:

Lot LL-1: Nominal 16 J energy specimens which exited the testing machine opposite the direction of swing of the pendulum.

Lot LL-44 & 33: Nominal 16 J energy specimens which exited the testing machine in the direction of swing of the pendulum (note: the material used for the tests employing the 2 mm striker were not from the same lot of material as the tests employing the 8 mm striker).

Lot HH-1: Nominal 100 J energy specimens which exited the machine in the direction of swing of the pendulum.

Five different test conditions were used. The five conditions were:

Condition 1: Rough anvils and a rough 8 mm radius striker.

Condition 2: Smooth anvils and a rough 8 mm radius striker.

Condition 3: Smooth anvils and a smooth 8 mm radius striker.

Condition 4: Smooth anvils and a rough 2 mm radius striker.

Condition 5: Smooth anvils and a smooth 2 mm radius striker.

Under Condition 1, five LL-1 specimens, five LL-44 specimens, and five HH-1 specimens were broken. The anvils had an rms surface finish of approximately 0.25 μm (10 μin.). The 8 mm striking bit had an rms surface finish of approximately 0.25 μm (10 μin.) on the nose and the 30° sides. On the parallel sides the rms surface finish was approximately 0.625 μm (25 μin.) (**Fig. 1**).

After the first tests, the anvils were removed and polished to a mirror finish. The rms surface finish measured 0.05 μm (2 μin.) after polishing. The anvils were reinstalled in the impact testing machine. Five LL-1 specimens, five LL-44 specimens, and five HH-1 specimens were then broken. This was Condition 2.

The striking bit was then removed and polished to a mirror finish. The nose, the 30° sides and the parallel sides were all polished to the same rms finish, approximately 0.05 μm (2 μin.).

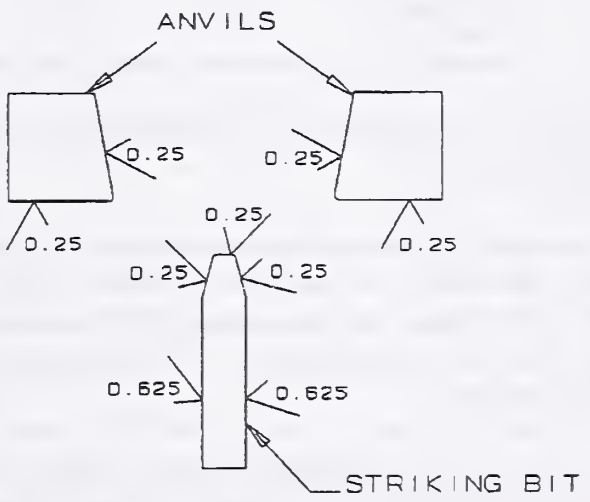


Figure 1. Surface finish (μm) of unpolished anvils and striking bit.

Table 1. Results of tests on specimens (lot LL-1) with a nominal energy of 16 J which exit the testing machine opposite the direction of swing of the pendulum. Test results are in joules.

Condition	1	2	3	4	5
	16.3	16.6	16.9	15.9	16.9
	16.9	16.6	16.6	16.6	16.9
	16.9	16.9	16.3	16.9	16.6
	16.5	15.9	15.6	18.0	16.3
	16.0	16.3	16.6	16.6	15.9
Average	16.52	16.46	16.40	16.80	16.52
Standard Deviation	0.35	0.34	0.44	0.68	0.38

Table 1A. Results of tests on specimens (lot LL-1) with a nominal energy of 16 J which exit the testing machine opposite the direction of swing of the pendulum. Test results are in joules. Outlier removed.

Condition	1	2	3	4	5
	16.3	16.6	16.9	15.9	16.9
	16.9	16.6	16.6	16.6	16.9
	16.9	16.9	16.3	16.9	16.6
	16.5	15.9	15.6	...	16.3
	16.0	16.3	16.6	16.6	15.9
Average	16.52	16.46	16.40	16.50	16.52
Standard Deviation	0.35	0.34	0.44	0.37	0.38

Five LL-1 specimens, five LL-44 specimens, and five HH-1 specimens were again broken. This was Condition 3.

The striking bit was again removed and a striking bit with a 2 mm radius nose was installed. The 2 mm striking bit had an rms surface finish of approximately 0.25 μm (10 μin.) on the nose and the 30° sides. On the parallel sides the rms surface finish was approximately 0.625 μm (25 μin.). Five LL-1 specimens, five LL-33 specimens, and five HH-1 specimens were broken. This was Condition 4.

The 2 mm striking bit was then removed and the nose, the 30° sides and the parallel sides were all polished to an rms finish of approximately 0.05 μm (2 μin.). Five LL-1 specimens, five LL-33 specimens, and five HH-1 specimens were again broken. This was Condition 5.

After a quick examination of the data, we decided to test five more LL-44 specimens which had been reserved. We removed the 2 mm striker and reinstalled the polished 8 mm striker. The five specimens were then tested under Condition 3.

Results

Twenty-five specimens were tested from lot LL-1 (**Table 1**). These specimens tended to exit the testing machine in a direction opposite the direction of swing of the pendulum. As seen in **Table 1**, there was very little difference in the average of the five different test conditions.

The condition with the highest variation was Condition 4 (smooth anvils and a rough 2 mm radius striking bit). An examination of the data for that group of tests indicated that the fourth specimen tested under Condition 4 was almost two standard deviations higher than the average. If this specimen is eliminated as an outlier, **Table 1A** is the result.

An examination of the averages and standard deviations listed in **Table 1A** yields very close agreement between the various conditions. The differences in the averages are much less than the standard deviations.

Table 2. Results of tests with an 8 mm striker on specimens (lot LL-44) with a nominal energy of 16 J which exit the testing machine in the direction of swing of the pendulum. Test results are in joules.

Condition	1	2	3	3
	17.5	17.6	17.3	16.3
	18.0	18.8	17.6	17.6
	17.3	17.4	16.9	17.3
	16.9	17.6	16.6	17.3
	18.0	17.4	17.3	17.6
Average	17.54	17.76	17.14	17.22
Standard Deviation	0.42	0.53	0.35	0.48
Average of 10	17.65		17.18	
Standard Deviation of 10	0.49		0.42	

Table 2A. Results of tests with an 8 mm striker on specimens (lot LL-44) with a nominal energy of 16 J which exit the testing machine in the direction of swing of the pendulum. Test results are in joules. Outliers removed.

Condition	1	2	3	3
	17.5	17.6	17.3	...
	18.0	...	17.6	17.6
	17.3	17.4	16.9	17.3
	16.9	17.6	16.6	17.3
	18.0	17.4	17.3	17.6
Average	17.54	17.50	17.14	17.45
Standard Deviation	0.42	0.12	0.35	0.17
Average of 9	17.52		17.28	
Standard Deviation of 9	0.32		0.34	

Twenty specimens were tested from lot LL-44 with an 8 mm striking bit (**Table 2**). These specimens tended to exit the testing machine in the direction of swing of the pendulum. Originally, five specimens were tested for each of the first three Conditions. After we noticed a

substantial drop in the absorbed energy between Conditions 2 and 3, five additional specimens were tested under Condition 3 so a greater population could be analyzed. There is a difference of approximately a 0.4 J (0.3 ft-lb) between the tests that were run before and after polishing the striking bit. This difference amounts to about 1 standard deviation.

Table 3. Results of tests with a 2 mm striker on specimens (lot LL-33) with a nominal energy of 16 J which exit the testing machine in the direction of swing of the pendulum. Test results are in joules.

Condition	4	5
	17.3**	16.6*
	16.9*	17.4**
	16.9**	16.9**
	16.9**	16.9
	17.3*	16.6**
Average	17.06	16.88
Standard Deviation	0.20	0.29

There is a difference of approximately a 0.2 J (0.15 ft-lb) between the tests done before and after polishing the anvils. The tests done with the polished anvils show a higher average absorbed energy than the average for tests done with unpolished anvils.

Two specimens in **Table 2** are almost two standard deviations from the average for their condition. **Table 2A** is a presentation of the results with these two specimens removed. Using this data, there is virtually no change between the results for polished and those for unpolished anvils. The difference between the polished and unpolished striking bit, although not as notable, is still evident.

Ten specimens were tested from lot LL-33 with a 2 mm striking bit (**Table 3**). We expected these specimens to exit the testing machine in the direction of swing of the pendulum. Many of the specimens exited in the opposite direction or did not exit the area of the anvils at all. Testing performed with an 8 mm striking bit led us to think that these specimens would exit the testing machine in the direction of swing of the pendulum. Perhaps it was a higher stress concentration due to the smaller radius striker used in these tests that caused the specimens in most cases to exit the machine in the direction opposite that of the pendulum or to not exit at all. Even though the specimens did not exit the machine in the direction of swing of the pendulum, a reduction in absorbed energy was observed when the striking bit was polished.

Twenty-five specimens were tested from lot HH-1 (**Table 4**). These specimens tended to exit the testing machine in the direction of swing of the pendulum. While the differences in the average absorbed energies are substantial by comparison with the energies for the low-energy specimens, with the exception of Condition 1 the difference between the test results is insignificant. The tests performed under Condition 1 with the unpolished anvils may have significantly higher energies than the other four cases with polished anvils.

Conclusion

For low-energy specimens which exit the testing machine in the direction opposite to the direction of swing of the pendulum, there seems to be very little difference between polished and unpolished anvils and strikers. There is only a slight downward trend in the average absorbed energies as the tooling is polished.

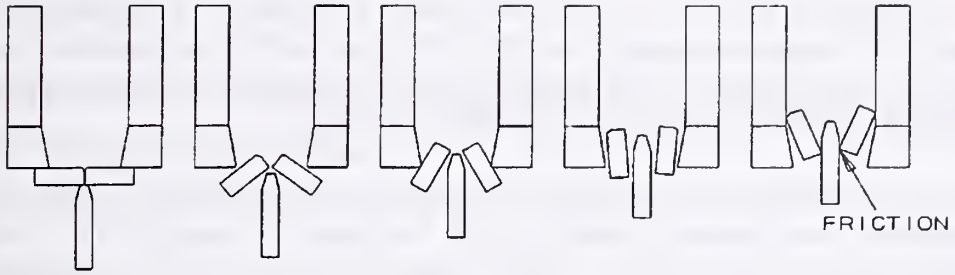


Figure 2. Proposed trajectories of low-energy specimens which cause lost energy due to friction with the striking bit.

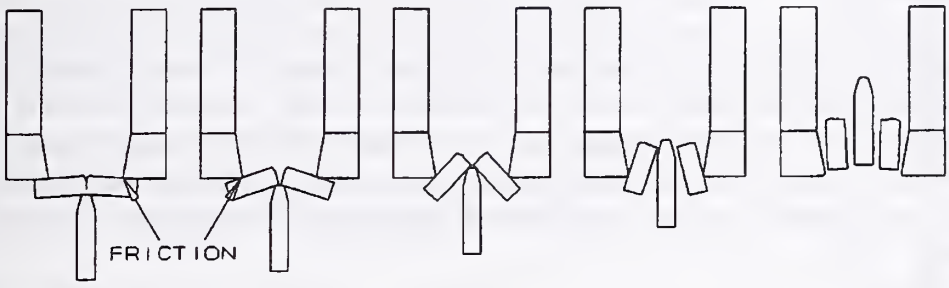


Figure 3. Proposed trajectories of high-energy specimens which cause lost energy due to friction with the anvils.

For low-energy specimens which exit the testing machine in the direction of swing of the pendulum, the average absorbed energy dropped as much as 0.4 J (0.3 ft-lb) when the striking bit was polished. When a low-energy specimen is broken, the broken halves spin at high velocity and strike the anvils a second time (**Fig. 2**). Specimens which exit the machine in the direction of swing of the pendulum rebound off the anvils and may continue to rotate in such a way as to scrape between the anvil support and the sides of the passing striking bit. We think that when we polish the sinking bit, the friction is reduced between the striking bit and the specimen after the specimen breaks.

Table 4. Results of tests on specimens (lot HH-1) with a nominal energy of 100 J which exit the testing machine in the direction of swing of the pendulum. Test results are in joules.

Condition	1	2	3	4	5
	86.8	89.2	96.6	95.6	92.5
	91.9	92.9	93.6	90.5	88.1
	92.2	85.8	87.1	85.1	91.9
	94.2	89.8	89.2	89.5	90.2
	92.2	89.8	86.1	94.2	91.2
Average	91.46	89.50	90.52	90.98	90.78
Standard Deviation	2.47	2.26	3.99	3.71	1.54

While the differences reported here may not seem large, 0.4 J can be important when compared to the ±1.4 J range permitted by ASTM E-23 when verification specimens are tested. A machine that may pass with an old striker that has been burnished by testing many specimens may fail when a new striker that is not highly polished is installed.

For high-energy specimens, we found that the absorbed energy dropped as much as 2 J (1.5 ft-lb) when the anvils were polished. Again a 1 or 2 J reduction in energy when attempting to qualify a machine can determine whether a machine passes or fails verification tests.

High-energy specimens are dragged across the anvils as they are broken (**Fig. 3**). Higher friction between the anvils and the specimen causes the testing machine to indicate higher absorbed energy values.

ASTM Standard Test Methods for Notched Bar Impact Testing of Metallic Materials (E 23) requires a 4 μm (125 μin.) rms finish on the anvils and the striker. Through use, these parts are burnished to a high polish. In order to reduce the shift in results that may result from a change in the striker or anvils, ASTM E 23 should require that these parts be highly polished. We suggest that ASTM E 23 be changed to require that the anvil and striking bit have rms finishes of 0.1 μm (4 μin.) or better.

References

[1] Siewert, T. A. and Vigliotti, D. P., "The Effect of Charpy V-notch Striker Radii on the Absorbed Energy," Pendulum Impact Machines: Procedures and Specimens for Verification. ASTM STP 1248.

C.N. McCowan and D.P. Vigliotti

The Influence of Shear Lip Symmetry on the Absorbed Energy of Charpy Impact Specimens

Introduction

It has been reported that the symmetry of the shear lip formation on Charpy impact specimens can have a significant influence on the absorbed energy.[1] For high energy level verification specimens, fracture surfaces having non-symmetric shear lips (**Figure 1**) have typically had higher absorbed energies than those having symmetric shear lips.

Here, data for six high energy pilot lots and two super high energy pilots lots are reported. The results are plotted in **Figures 2 to 25**, and summary data for the pilot lots are given in **Table 1** .

Result and Discussion

The data are grouped by pilot lot and test machine. This is because each pilot lot has a different absorbed energy, and there are differences between the three machines used for the testing of a given pilot lot.

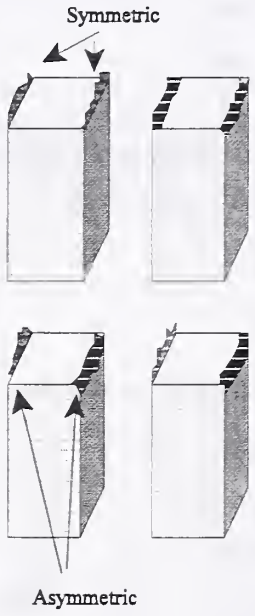


Figure 1: Diagram showing: (a) symmetric shear lips, which have both shear lips (up) on the same fracture surface, and (b) asymmetric shear lips, which have one shear lip up on each fracture surface.

The data consistently show a trend of higher energy for specimens that fractured asymmetrically. In **Figure 26**, the difference in the energy between the asymmetric and symmetric specimens ranges from near 0 to 5 J, for the high-energy specimens, and from about 5 to 12 J for the super-high energy specimens. Assuming the typical energy levels for the high and super-high energy specimens are about 100 and 200 J respectively, the maximum increase in energy is near 5 % for both types of specimens. This is significant, because the requirements for the impact verification test only allows ±5 % variation from the certified value of the specimen.

The effect of specimen squareness on the test results was also evaluated. No correlation between squareness and energy (or fracture symmetry) was found. However, we are not convinced that squareness is not an important variable, so more testing is planned.

Currently we have no quantitative method to evaluate the influence of fracture symmetry on the results of verification tests. However, we do consider the mix of fracture types in our customer results (and the mix from the original pilot lot data) when we evaluate the data for a given machine. If, for example, a customer result was slightly high and they failed the verification test for no apparent reason, and all five of the customers specimens had asymmetric fractures, we might have them retest.

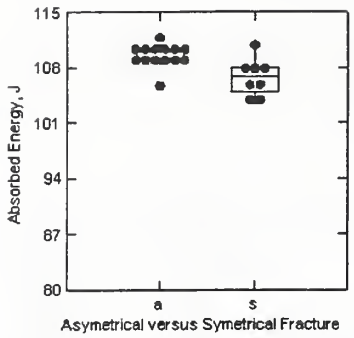


Figure 2: HH84 SI2

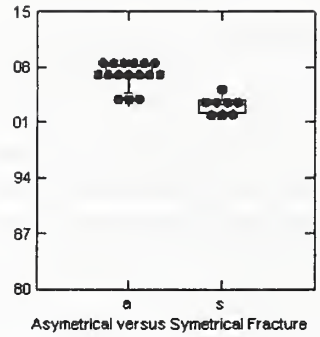


Figure 3: HH84 TK

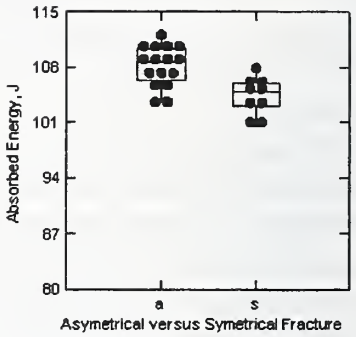


Figure 4: HH84 TO2

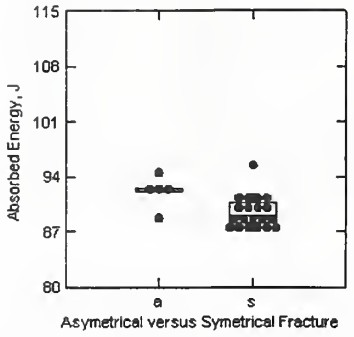


Figure 5: HH86 SI2

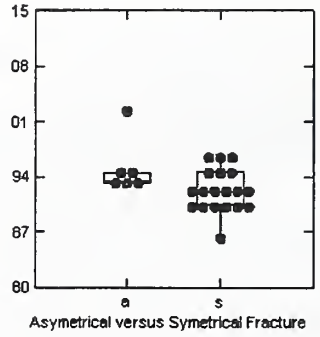


Figure 6: HH86 TK

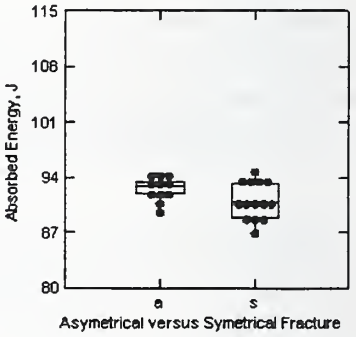


Figure 7: HH86 TO2

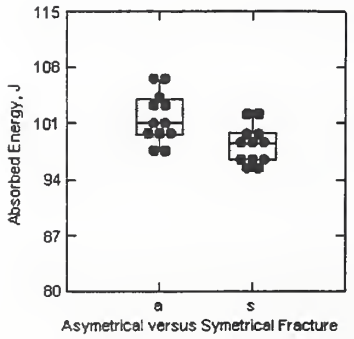


Figure 8: HH87 SI2

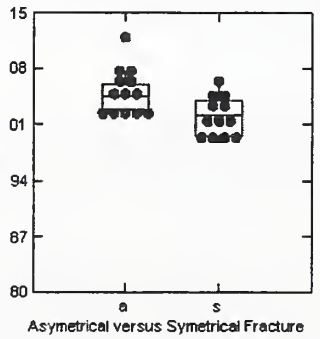


Figure 9: HH87 TK

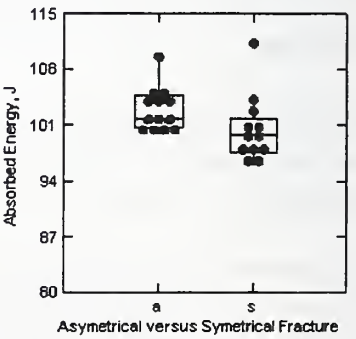


Figure 10: HH87 TO2

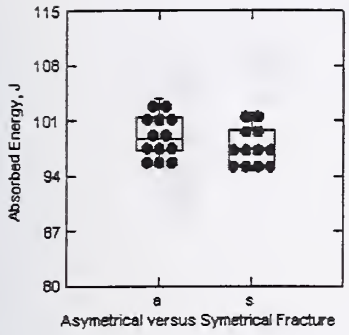


Figure 11: HH92 SI2

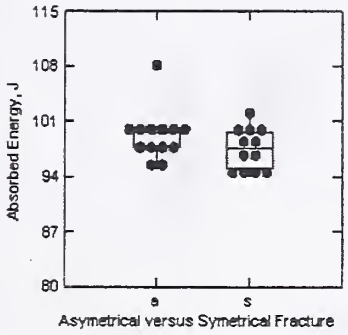


Figure 12: HH92 TK

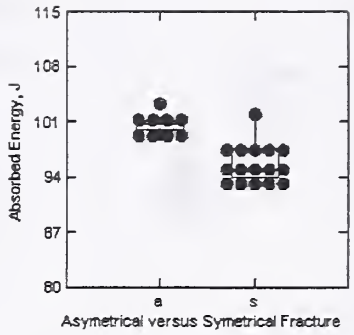


Figure 13: HH92 TO2

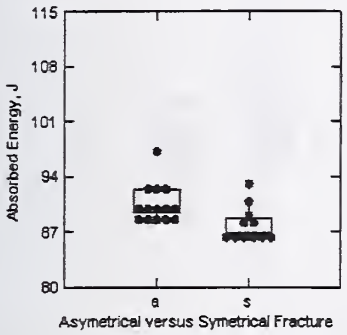


Figure 14: HH93 SI2

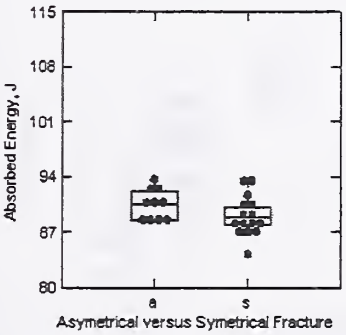


Figure 15: HH93 TK

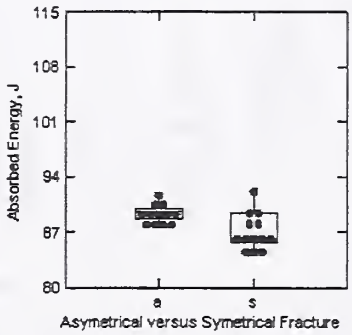


Figure 16: HH93 TO2

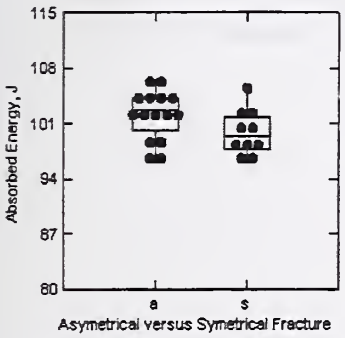


Figure 17: HH94 SI2

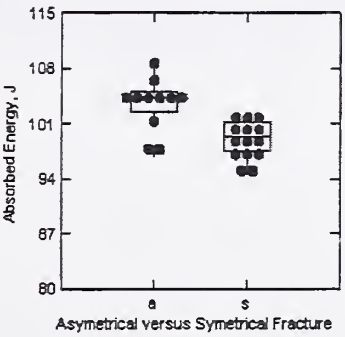


Figure 18: HH94 TK

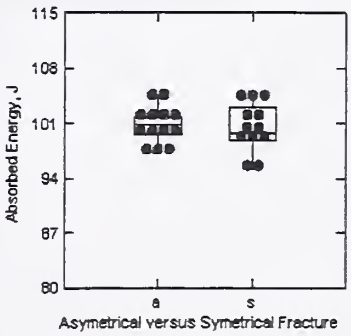


Figure 19: HH94 TO2

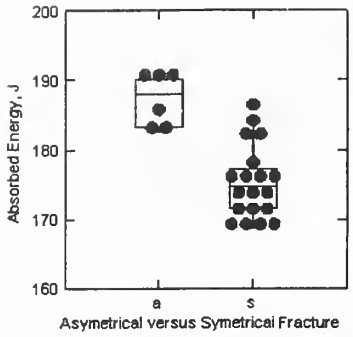


Figure 20: SH16 SI2

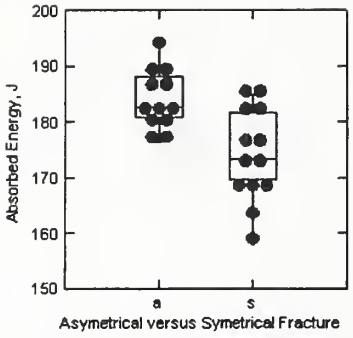


Figure 21: SH16 TK

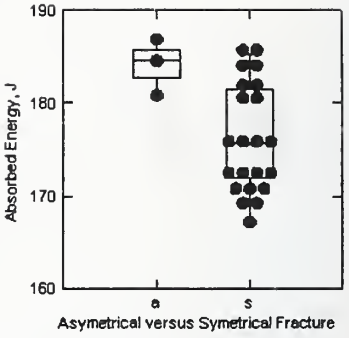


Figure 22: SH16 TO2

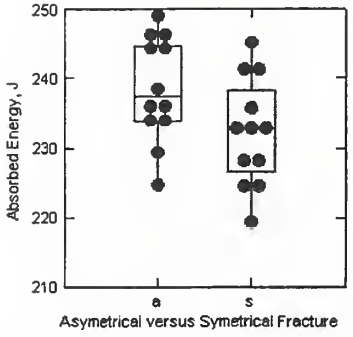


Figure 23: SH24 SI2

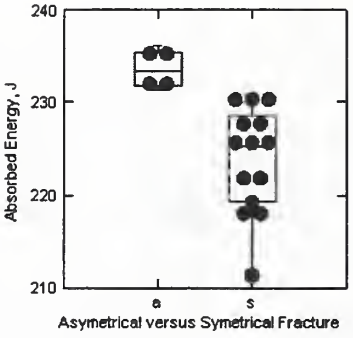


Figure 24: SH24 TK

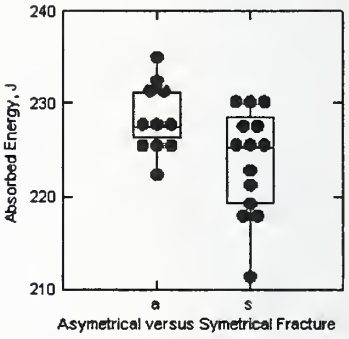


Figure 25: SH24 TO2

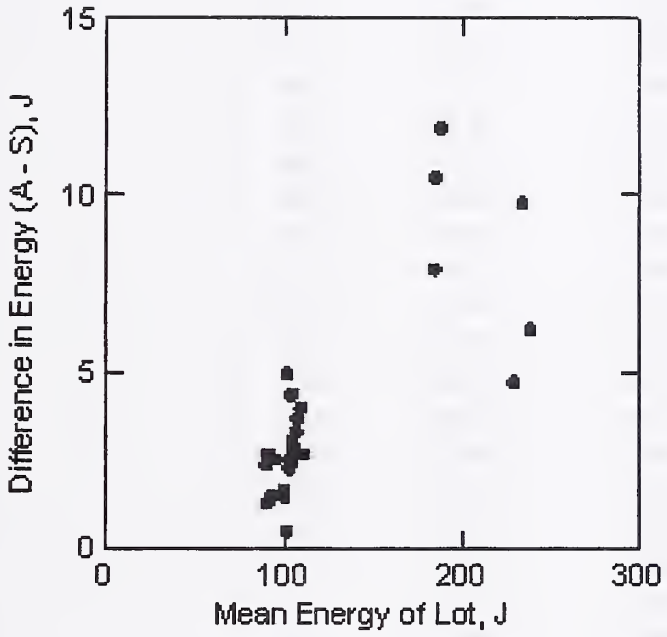


Figure 26: The difference in the energy of the asymmetric fractures and the symmetric fracture for high and super-high energy levels.

Table 1. Summary statistics for pilot lots.

Pilot Lot SERIES	MACHINE	Asymmetrical Energy, J	Symmetrical Energy, J	Difference J
HH93	TO2	89.4	87.1	2.4
HH93	TK	90.4	89.2	1.3
HH93	SI2	90.7	88.0	2.7
HH86	SI2	92.1	89.5	2.6
HH86	TO2	92.5	91.0	1.5
HH86	TK	94.9	92.4	2.5
HH92	SI2	99.1	97.7	1.5
HH92	TK	99.3	97.6	1.6
HH92	TO2	100.6	95.7	4.9
HH94	TO2	100.8	100.4	0.5
HH94	SI2	102.0	99.8	2.3
HH87	SI2	102.8	98.5	4.3
HH87	TO2	102.9	100.5	2.4
HH94	TK	103.5	99.1	4.4
HH90	SI2	103.5	100.9	2.6
HH89	TO2	103.9	100.8	3.1
HH87	TK	104.9	102.0	2.9
HH88	TO2	105.5	102.3	3.3
HH88	SI2	106.8	103.2	3.6
HH84	TK	106.8	103.1	3.7
HH84	TO2	108.3	104.4	4.0
HH84	SI2	109.4	106.7	2.7
SH16	TO2	184.1	176.2	7.9
SH16	TK	184.3	173.8	10.5
SH16	SI2	187.1	175.2	11.9
SH24	TO2	228.5	223.8	4.7
SH24	TK	233.6	223.8	9.8
SH24	SI2	238.4	232.2	6.2

Reference

1. C.N. McCowan, J. Pauwels, G. Revice, and H. Nakano, International Comparision of Impact Verification Programs, in Pendulum Impact Testing: A Century of Progress, ASTM STP1380, p. 73, 1999.

Part 4: Statistical Evaluations of Charpy Impact Data

C.N. McCowan,[1] C.M. Wang,[2] and D.P. Vigliotti[1]

Charpy Impact Verification Data (1994-1996): A Summary

Reference: McCowan, C. N., Wang, C. M., and Vigliotti, D. P., "**Charpy Impact Verification Data (1994-1996): A Summary,**" *Journal of Testing and Evaluation*, JTEVA, Vol. 27, No. 2, March 1999, pp. 89-99.

Abstract: We present a summary of Charpy impact verification test data that were evaluated by the National Institute of Standards and Technology from January 1994 to December 1996. The Charpy impact machines that met the verification requirements of ASTM Test Methods for Notched Bar Impact Testing of Metallic Materials (E 23) are broken down by year and by reference lot. Based on the data, a proposed verification rule that limits the range of the verification set has been examined. We also present the results for determining whether two energies (lower and upper ends of the machine capacity) or three (lower, middle, and upper) are needed to verify the performance of the large-capacity impact machines; currently E 23 requires three energies to be tested.

Keywords: Charpy V-notch, impact certification program, impact testing, notched-bar testing, pendulum impact machines, reference specimens

This report provides a summary of the Charpy impact verification data that were evaluated by the National Institute of Standards and Technology (NIST) from January 1994 through December 1996. An indirect verification program has been used to verify the performance of Charpy impact machines for more than 40 years [1,2], and in 1964 ASTM Standard E 23 [3] was revised to require verification testing. NIST has provided the verification specimens and administered the program since 1989. In this program, impact machines are verified annually to the requirements in E 23, and the verification data, which are generated by organizations that own test machines, are returned to NIST for evaluation.

The impact verification program can be divided into three basic parts: production and distribution of impact verification specimens; verification testing; and evaluation of the verification test data. Before verification testing, a reference value for the impact toughness of the verification specimens is determined, and the uncertainty associated with the reference value is confirmed to be below a limit that ensures the material homogeneity of the specimens. In practice, the reference value for the impact specimens is determined by testing a random sample of 75 specimens from a production lot, which normally contains 1200 heat-treated specimens.

The 75 specimens are divided into groups of 25 and tested on three machines that have been defined in E 23 as the reference impact machines for the United States. The impact toughness, defined as the energy absorbed in the test, is the average absorbed energy for the75 tests. If the

1National Institute of Standards and Technology, Materials Reliability Division, Boulder, CO 80305
[2] National Institute of Standards and Technology, Statistical Engineering Division, Boulder, CO 80305

lot meets the statistical requirements of ASTM Practice for Qualifying Charpy Verification Specimens of Heat-Treated Steel (E 1271), a reference value is assigned to the lot [4]. Once a lot is accepted, sets of five specimens, or verification sets, are sold to companies that want to verify the performance of their Charpy impact machines. The specimens are broken using the candidate machine, and the broken specimens, along with the absorbed energy results, are sent to NIST for analysis. If the results are within 5 or 1.4 J of the reference value and the markings left by the machine on the specimens indicate the machine is in good working condition, the candidate machine is certified by NIST to meet the requirements of E 23.

A database containing the results of verification tests has been collected that includes the serial number of the candidate machine, the capacity and the pendulum design of the machine, the energy obtained for each specimen tested, the reference energy for the specimens tested, and the date of the test. The principal use of these data is to track the performance of individual impact machines and to monitor the verification program. The data also provide an opportunity for cross-validation of the current acceptance criteria and evaluation of new criteria proposed for the verification of Charpy impact machines.

Data from the verification tests and pilot lot evaluations are presented here to provide a general overview of how the Charpy verification program works in practice. To do this, tables and graphics are employed to show the proportions of test results that meet the current verification requirements and to consider the influence of the verification specimen on the test results. The data are also used to examine a newly proposed range rule for verification tests. Finally, the data are used to evaluate the requirement in E 23 to test at three energies (lower, middle, and upper ends of the machine capacity) when verifying large impact machines over their full capacity. The three energies available for verification testing are referred to as low, high, and super-high energy specimens in this paper. There is an interest in determining whether only two energies (low and super-high) are sufficient to verify the performance of the large-capacity impact machines.

Materials and Procedures

Charpy impact verification specimens are sold by NIST at three energies: the low-energy specimens, with energies near 17 J, the high energy specimens, with energies near 100 J, and the super-high energy specimens with energies near 225 J. The low- and high-energy specimens are made from 4340 steel, which is heat treated to produce specimens at the appropriate absorbed energies. The super-high-energy specimens are made using a T-200 maraging steel. The data used here include the customer data from many different lots of low-, high-, and super-high-energy specimens. The total number of verification tests for the low-, high-, and super-high-energy levels are 2401, 2385, and 655, respectively.

If the five absorbed energy measurements from the Charpy machine being tested are denoted by e_1, e_2, e_3, e_4, e_5, and E_c is the average; of e_i, i = 1, 2, ···, 5, then the low-energy certification criterion is given by:

$$-1.4 \leq E_c - E_r \leq 1.4$$

where E_r is the reference energy of the lot. To certify a candidate Charpy machine in the high- or super-high-energy ranges, the difference between E_c and E_r must be within 5 % of the reference value; that is, it requires:

$$-0.05 \leq (E_c - E_r)/E_r \leq 0.05$$

Results

Low Energy

Figure 1 plots the difference between the customer's average and the reference value, $E_c - E_r$, for the low-energy verification tests. The dotted lines indicate the 1.4 J pass or fail criteria. In the right margin, the density estimate for $E_c - E_r$ is also plotted. The distribution of $E_c - E_r$ is centered around 0 and has a longer "tail" on the positive side of $E_c - E_r$. Out of 2401 verification tests shown in **Fig. 1**, 283 tests (11.8 %) failed the indirect verification. **Figure 1** also shows that when a machine fails to meet the low-energy verification requirements, it fails more often with $E_c - E_r > 1.4$ (9.2 %) than with $E_c - E_r < -1.4$ (2.6 %). This is expected because most of the common factors, such as anvil radius, bearings, and mounting, that wear or loosen over time increase the energy absorbed by the machine during the test.

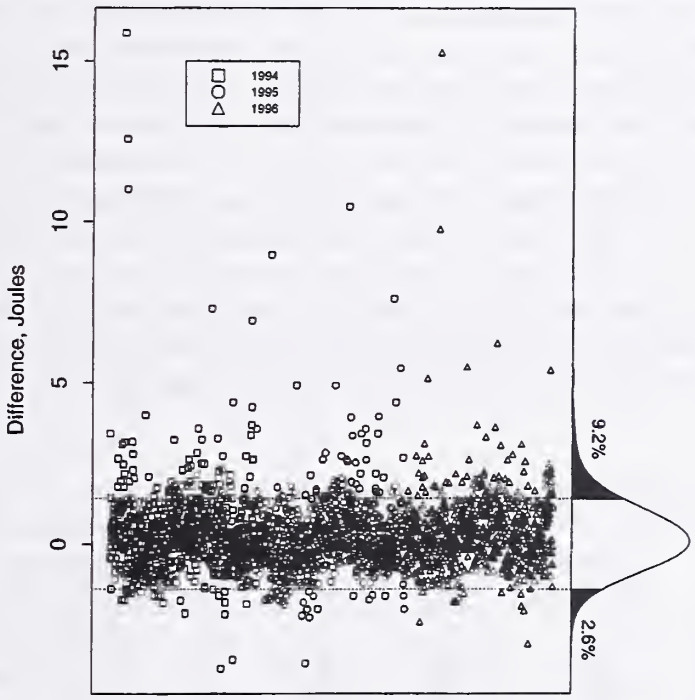


Figure 1. The difference, in J, between the customer's average for low energy verification tests and the reference value. The horizontal axis is the sequence of the test. The order within each year is not relevant. The dotted lines indicate the 1.4 J pass/fail criteria. The density estimate for the differences is plotted in the right margin.

Table 1. Pass/fail data for low energy verification tests (by year).

Year	Status	$E_c \leq E_r$	$E_c > E_r$	Total
1994	Pass	314 (39.3%)	388 (48.5%)	702 (87.8%)
	Fail	20 (2.5%)	78 (9.8%)	98 (12.2%)
1995	Pass	350 (43.6%)	359 (44.7%)	709 (88.3%)
	Fail	31 (3.9%)	63 (7.8%)	94 (11.7%)
1996	Pass	294 (36.8%)	413 (51.8%)	707 (88.6%)
	Fail	11 (1.4%)	80 (10.0%)	91 (11.4%)
Total	Pass	958 (39.9%)	1160 (48.3%)	2118 (88.2%)
	Fail	62 (2.6%)	221 (9.2%)	283 (11.8%)

In **Table 1**, the pass/fail data are broken down into test year. The number inside the parentheses is the percentage of the respective yearly total. The percentages of the row "Total" are with respect to the total number of the tests. The data show that the failure rate for the low-energy verification tests remains fairly constant (12.2 %, 11.7 %, and 11.4 %) from year to year.

In **Table 2**, the pass/fail data are broken down by the lot designation of verification specimens tested. The first column is the lot identification; the second column is the reference value of absorbed energy (J) of the lot; the third column is the standard deviation (J) of the pilot lot (calculated as the square root of the weighted mean of the variances of the three machines with weights equal to the number of observations, i.e., the pooled estimate); the fourth column is the lot size, or the number of machines tested using the verification specimens from that lot; the next two columns are, respectively, the number of machines that passed the verification test on the low and high sides; and the last two columns are the number of machines that failed the test and the failure rate. Only lots for which at least 100 verification tests were made are listed here. We present the specimen lot data in this way to evaluate the influence of the specimens on the outcome of the verification tests. The standard deviation of the pilot lot is one of the primary subtle trend for increasing failure over the small range in standard deviation present in the data. Unlike the yearly data, where the same population of machines is compared (since machines are tested annually), the failure rates of lots are expected to vary because the number of machines

Table 2. Pass/fail data for low energy verification tests (by lot).

Lot	Reference Energy	Standard Deviation	Lot Size	No. of Passes		No. of Failures	Failure Rate
				Low	High		
44	17.2	0.62	132	46	70	16	12.1%
45	17.4	0.52	167	62	90	15	9.0%
46	17.4	0.58	165	90	69	6	3.6%
47	17.2	0.50	174	90	61	23	13.2%
48	17.6	0.72	162	79	67	16	9.9%
49	17.6	0.65	172	59	90	23	13.4%
51	16.9	0.71	148	61	64	23	15.5%
52	17.2	0.71	129	63	54	12	9.3%
53	16.7	0.57	130	49	70	11	8.5%
54	16.8	0.65	114	16	72	26	22.8%
55	17.5	0.84	117	48	54	15	12.8%
56	18.2	0.73	108	33	65	10	9.3%

tested for each lot is some different fraction of the machine population tested each year. So the acceptance criteria for the lots, and lower variation should result in less influence by the specimens on the outcome of the test. The fact that there is not a trend of increasing failures with increasing standard deviation in **Table 2** is not surprising, however, because much of the variation in the failure rates of the lots is probably due to sampling, which would obscure any standard deviation cannot be related to the failure rate for these data, and the failure rate alone should not be considered to have too much significance.

Probably the best indicator of specimen influence on the verification test we have is the distribution of the pass (low or high) data shown in **Table 2**. These data, monitored periodically as they are accumulated on each lot, show how many test results were higher or lower than the reference value assigned to the verification specimens being tested. A very skewed result here indicates that the reference energy assigned to the verification specimens could be very different from the average of the test results for the *good* machines. We would not expect the two averages to be identical, but if the *good* machines tested are a representative sample of the population of *good* machines, we typically find good agreement between the two averages. So, when the pass data are very skewed and the failure rate is high, as for Lot 54, we suspect the lot might be influencing the verification test results.

To more fully evaluate the data for Lot 54, the distribution in energy for the customer data is compared to the distribution of the pilot lot data in **Fig. 2**. It shows that the distribution of the customer data for Lot 54 is approximately Gaussian, with a slightly higher average energy than the pilot lot data (the average energy of the pass data for Lot 54 is 17.2 J). The peak around 18 J in the customer data is not considered to be part of the distribution of good machines. Rather, it defines a population of machines that failed the test and that differs from the population of *good* machines (both pass and fail data were included in the plot of the customer data). Although the apparent shift in average energy for the customer data may push some good machines across the pass/fail limit, most of the machines that failed appear to be representative of a population consisting of bad machines. Note also the almost bimodal shape near

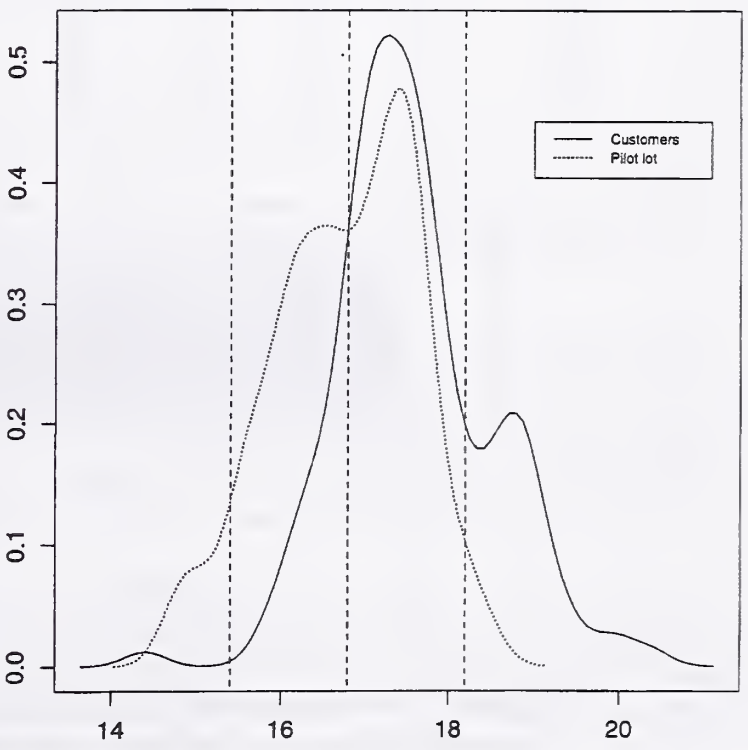


Figure 2. The density plots of the energy absorbed for the pilot lot data (from 75 individual specimens) associated with Lot 54. The three vertical dashed lines correspond to the reference energy (16.8 J) and the acceptance region (16.8 ± 1.4 J).

the peak of the distribution of the pilot lot data. This example appears somewhat extreme, but is not unexpected, because the pilot lot data are a compilation of data from three machines, each of which has characteristic differences and biases compared with the two others. For this pilot lot, two machines had very similar average energies (17.3 and 17.0 J), which differed from the average of the third machine (16.0 J). The combined data of the three reference machines serve as a good illustration of how the inclusion of different machines results in a balanced average reference energy for the specimens, but the distribution is broadened due to the bias between machines and does not provide a good measure of the inherent scatter of the specimens. To remove the machine bias and better estimate the scatter of the specimens, the pooled standard deviation of the three machines is used. For this pilot lot, the pooled standard deviation was 0.65 J, while the (not pooled) standard deviation for the (combined) data shown in **Fig. 2** was 0.87 J.

High Energy

Figure 3 plots the relative difference between the customer's average and the reference value, $(E_c - E_r)/E_r$, for the high-energy verification tests. The dotted lines indicate the 5 pass or fail criteria. In the right margin, the density estimate for $(E_c - E_r)/E_r$ is also plotted. The distribution of $(E_c - E_r)/E_r$ is centered around 0 and has a slightly longer "tail" on the positive side of $(E_c -$

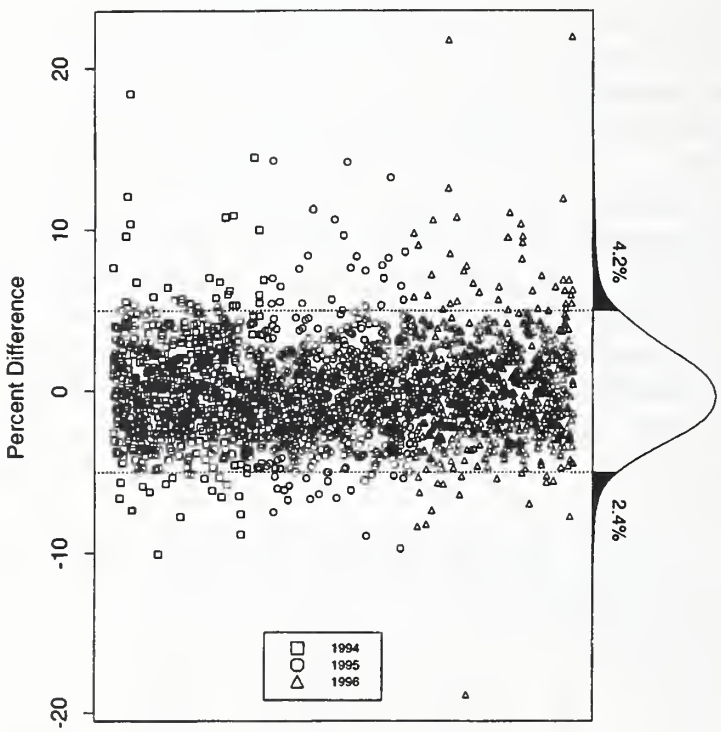


Figure 3. The relative difference between the customer's average for high-energy verification tests and the reference value. The horizontal axis is the sequence of the test. The order within each year is not relevant. The dotted lines indicate the 5% pass/fail criteria. The density estimate for the relative differences is plotted in the right margin.

Table 3. Pass/fail data for high-energy verification tests (by year).

Year	Status	$E_c \leq E_r$	$E_c > E_r$	Total
1994	Pass	363 (46.8%)	357 (46.1%)	720 (92.9%)
	Fail	21 (2.7%)	34 (4.4%)	55 (7.1%)
1995	Pass	418 (53.9%)	313 (40.3%)	731 (94.2%)
	Fail	19 (2.4%)	26 (3.3%)	45 (5.8%)
1996	Pass	397 (47.6%)	379 (45.4%)	776 (93.0%)
	Fail	17 (2.0%)	41 (4.9%)	58 (7.0%)
Total	Pass	1178 (49.4%)	1049 (44.0%)	2227 (93.4%)
	Fail	57 (2.4%)	101 (4.2%)	158 (6.6%)

Table 4. Pass/fail data for high energy verification tests (by lot).

Lot	Reference Energy	Standard Deviation	Lot Size	No. of Passes		No. of Failures	Failure Rate
				Low	High		
44	98.2	2.48	106	51	54	1	0.9%
45	99.1	2.62	160	76	79	5	3.1%
46	100.7	2.94	167	78	70	19	11.4%
47	108.6	3.42	169	125	19	25	14.8%
48	103.1	3.25	182	94	77	11	6.0%
49	102.6	3.16	183	65	106	12	6.6%
51	101.1	3.09	191	74	109	8	4.2%
52	102.2	2.96	123	49	70	4	3.3%
53	98.2	2.71	136	28	98	10	7.4%
54	99.8	2.70	136	70	62	4	2.9%
55	97.5	3.47	112	62	39	11	9.8%
56	106.7	3.58	121	65	41	15	12.4%

$E_r)/E_r$. Out of 2385 verification tests shown in **Fig. 3**, 158 tests (6.6 %) fail to pass the indirect verification requirements, with 4.2 % failing on the high side and 2.4 % on the low side.

In **Table 3**, the pass/fail data are broken down into test year. Like the low energy data, the failure rate for the high-energy verification tests remains fairly constant (7.1 %, 5.8 %, and 7.0 %) from year to year. The failure rate for the high-energy verification tests is consistently lower than that for the low-energy verification tests. It has long been recognized by E 23 that the very-high-strength, low-energy impact specimens show performance problems with machines that high-energy specimens do not. The difference in pass/fail percentages highlights this point.

Table 4 displays the pass/fail data by lot of verification specimens. Again, with the small number of machines tested under each lot, the failure rates vary from lot to lot. Considering the first three lots (44, 45, and 46), for example, the standard deviations and distributions of the pass data are similar, but the failure rates vary greatly. We assume that the explanation here is that a higher percentage of *bad* machines were tested using the Lot 46 specimens. However, for Lot 47, which has a very skewed pass data, we cannot necessarily attribute the high failure rate to sampling.

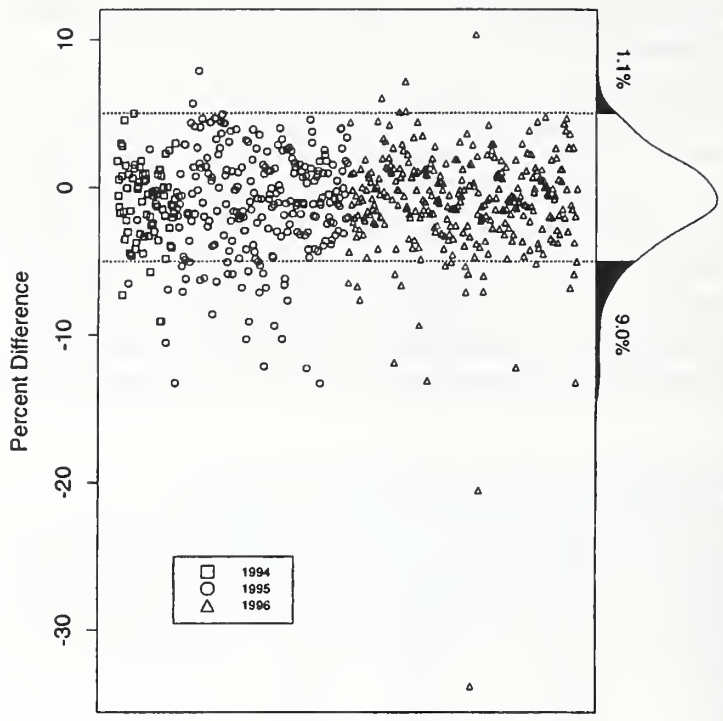


Figure 4. The relative difference between the customer's average for super-high-energy verification tests and the reference value. The horizontal axis is the sequence of the test. The order within each year is not relevant. The dotted lines indicate the 5% pass/fail criteria. The density estimate for the relative difference is plotted in the right margin.

Super-High Energy

Figure 4 plots the relative difference between the customer's average and the reference value, $(E_c - E_r)/E_r$, for the super-high energy verification tests. The dotted lines indicate the 5 % pass or fail criteria. In the right margin, the density estimate for $(E_c - E_r)/E_r$ is also plotted. The distribution of $(E_c - E_r)/E_r$ has a longer "tail" on the negative side of$(E_c - E_r)/E_r$. This implies when a machine fails the super-high verification test, it tends to have a low E_c value. This differs from that of the low- and high-energy verification data and is suspected to be a result of the different specimen-anvil (and striker) interaction for the super-high-energy test. These very ductile specimens are deeply brinelled by the anvils and wrap around the striker during the test. Overall, 10.1 % of the 652 cases evaluated failed to meet the 5 % verification criteria of E 23.

Table 5 contains the pass/fail data broken down into test year for the super-high verification tests. The failure rate varies more for these data than it does for the high- or low-energy verification data, but this is a new test. The number of machines tested from year to year is not as constant as in the low- or high-energy tests, and many machines are being tested for the first

Table 5. Pass/fail data for super-high-energy verification tests (by year).

Year	Status	$E_c \leq E_r$	$E_c > E_r$	Total
1994	Pass	47 (56.0%)	30 (35.7%)	77 (91.7%)
	Fail	7 (8.3%)	0 (0.0%)	7 (8.3%)
1995	Pass	117 (47.8%)	97 (39.6%)	214 (87.3%)
	Fail	29 (11.8%)	2 (0.8%)	31 (12.7%)
1996	Pass	182 (55.8%)	116 (35.6%)	298 (91.4%)
	Fail	23 (7.1%)	5 (1.5%)	28 (8.6%)
Total	Pass	346 (52.8%)	243 (37.1%)	589 (89.9%)
	Fail	59 (9.0%)	7 (1.1%)	66 (10.1%)

Table 6. Pass/fail data for super-high energy verification tests (by lot).

Lot	Reference Energy	Standard Deviation	Lot Size	No. of Passes		No. of Failures	Failure Rate
				Low	High		
3	229.9	7.25	135	81	33	21	15.6%
4	226.3	6.78	159	73	74	12	7.6%
5	222.5	8.12	169	81	77	11	6.5%
6	224.8	5.59	158	92	52	14	8.9%

time at these energies. For these reasons, we anticipate that the 10.1 average failure rate for the super-high energy tests will decrease slightly in the years to come.

Table 6 displays the pass/fail data by lot of verification specimens. Lot 3 has the highest failure rate among the lots for which we now have data, but we suspect that many of the failures associated with this first lot were due to not testing the specimens at room temperature (as required). The rest of the lots have comparable failure rates.

Range Summary

The range of the five measurements in a verification test is defined as

$$R = \max\{e_i\} - \min\{e_i\}$$

A new verification rule limiting the range of the absorbed energy measurements for the low- and high-energy tests is being balloted for ASTM E 23. The range test is designed to detect excessive variation, that is, to identify machines that have very high scatter in their measurements and just happen to have mean energy values that agree with the reference energy. Splett and Wang [5] also proposed an alternate certification procedure that accounts for the lot and machine variations.

The ranges being considered for limiting the low- and high-energy tests are 5 and 15 J, respectively. So, for the low-energy test, a candidate machine would fail if its range were greater

than 5 J even it passed the ±1.4 J criteria, and for the high-energy test, a candidate machine would fail if its range were greater than 15 J even it passed the ±5% criteria.

Based on the data from 1994 to 1996, these range limits appear reasonable **Figs. 5 and 6** display the range of the five measurements for the low- and high-energy verification tests. The solid markers designate the machines that fail the ± 1.4 J criteria in the low-energy test (**Fig. 5**) and ±5% criteria in the high energy test (**Fig. 6**). There are 16 tests with $R > 5$ in **Fig. 5** and 17 tests with $R > 15$ in **Fig. 6**. If the range rule were in use, nine additional tests would fail in the low energy test, increasing the failure rate from 11.8 % to 12.2 %. Similarly, eight additional tests would fail in the high-energy test, increasing the failure rate from 6.6 to 7.0 %. In both cases, the range rules would increase the failure rate by 0.4 %.

An alternate range rule would be to use the normalized range

$$R_n = R/E_r$$

and fail the candidate machine if $R_n > r$, where r is some specified limit. The range rule based on R_n enables us to have the possibility of using one limit for all the three energies. It also provides a useful interpretation for the rule. It can be shown (e.g., see Ref 6) that for samples of five observations from a Gaussian distribution

$$R \approx 2.33\, S$$

where S is the standard deviation. Thus,

$$R_n \approx 2.33\, S/E_r$$

and the rejection criterion $R_n > r$ is approximately equivalent to the rejection criterion

$$S/E_r > r/2.33$$

which can be interpreted to mean that (in additional to the regular "difference" criteria) a machine would fail the test if its noise-to-signal ratio is greater than $r/2.33$. For example, with $r = 25$, the threshold would be 10.73 %.

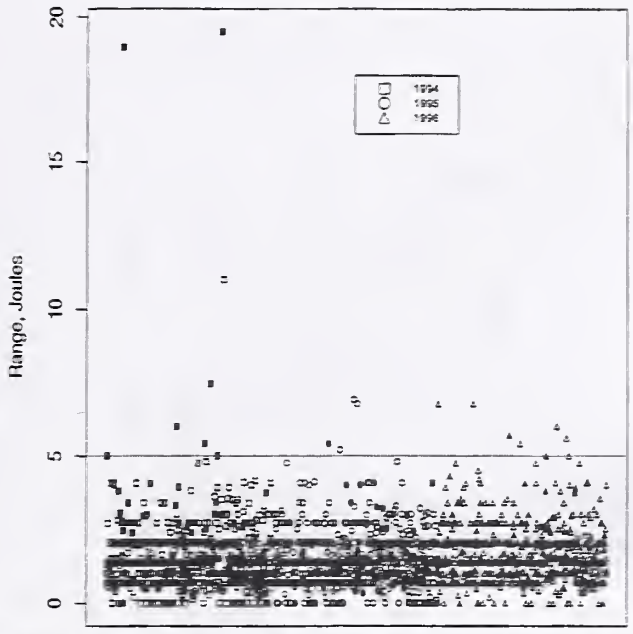


Figure 5. The range of the five specimens for the low-energy verification test. The horizontal axis is the sequence of the test. The order within each year is not relevant. The dotted line indicates the 5 J proposed rule. The solid markers designate the machines that fail the ±1.4 J criteria.

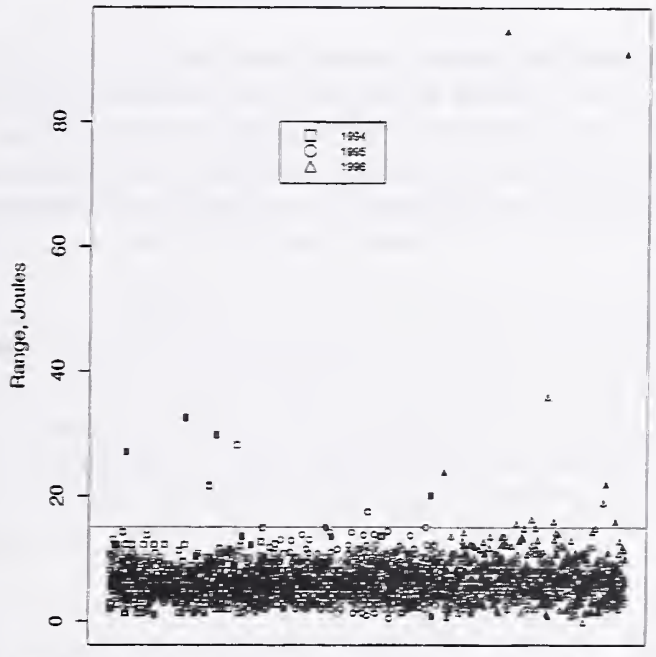


Figure 6. The range of the five specimens for the high-energy verification test. The horizontal axis is the sequence of the test. The order within each year is not relevant. The dotted line indicates the 15 J proposed rule. The solid markers designate the machines that fail the ±5% criteria.

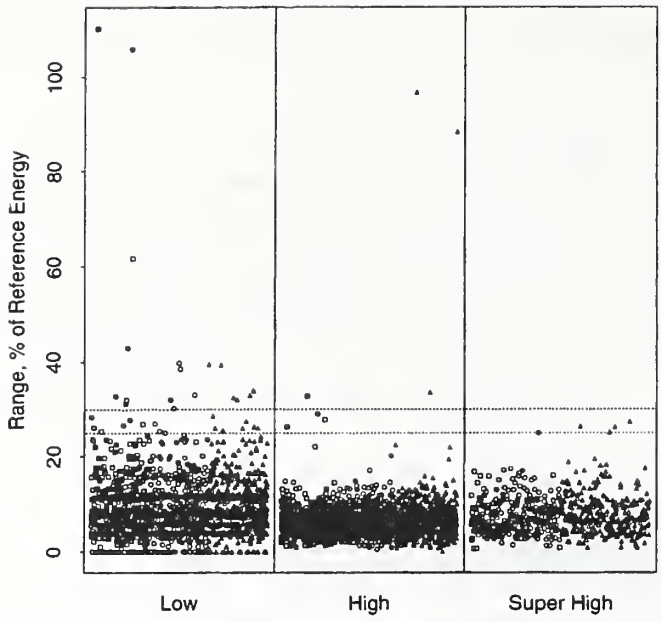


Figure 7. The normalized range of the five specimens for all the three energy verification tests. The solid markers designate the machines that fail the ±1.4 J or ±5% criteria. The dotted lines, at 25% and 30 %, are possible values of r to use in the range value.

Figure 7 plots R_n for all three energy verification tests. Again, the solid markers designate the machines that fail the ±1.4 J or ±5 % criteria. The two dotted lines indicate the two possible values of r to use: 25 and 30 %, which correspond to 10.73 and 12.88 % in the noise-to-signal-ratio scales. A 30 % R_n criterion would fail eleven additional tests (0.46 %) for the low energy, one additional test (0.04 %) for the high, and no additional test for the super-high energy.

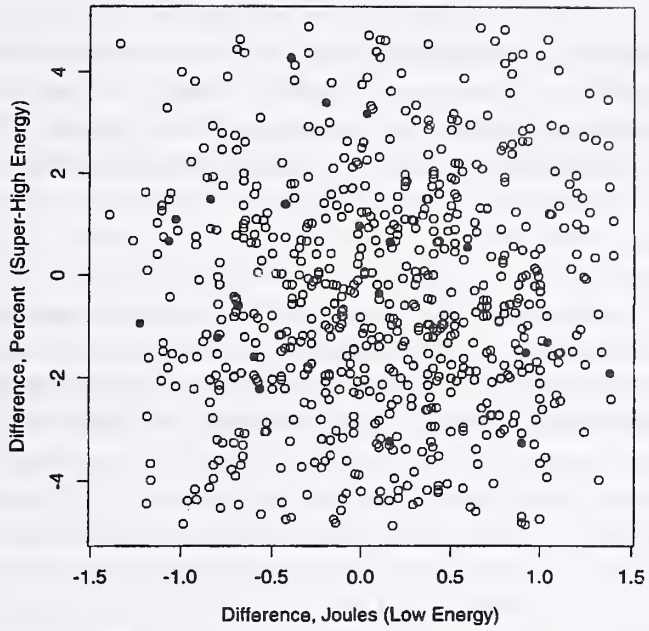


Figure 8. The scores of the low and super-high energy tests for cases that passed both tests. The "○" points are cases that also passes the high energy tests.

Energies Required for Verification Testing

Currently, the testing of the low, high, and super-high energies is required by ASTM Standard E 23 to certify candidate machines with capacities of greater than 289 J. However, the responsible ASTM subcommittee and task group has questioned whether the testing of the high energy is necessary in the certification of high-capacity machines. To address this question, we examine the verification tests from 1994 to 1996 for which the low-, high-, and super-high-energy tests were performed. **Figure 8** plots the values of $E_c - E_r$ (in the low-energy test) and $(E_c - E_r)/E_r$ (in the super-high test) for the cases that passed the low- and super-high-energy tests. The 698 "o" points are cases that also passed the high energy test. The 21 "●" points (2.92 %) are cases that failed the high energy test. Among the 21 cases that failed, 13 tests failed with $(E_c - E_r)/E_r$ between 5 and 6 %, and five tests failed with $(E_c - E_r)/E_r$ between 6 and 7 %. Thus, the successful result of the low- and super-high energy verification tests cannot ensure that a machine will pass the high-energy verification criteria of ±5 %, but in most cases, the results of the high-energy test are in good agreement with the reference energy.

Discussion

The failure rates for the low and high-energy verification tests are nearly constant, averaging 11.8 and 6.6 %, respectively. The failure rate for the super-high-energy test is more variable, in part due to the smaller and unequal number of tests performed in each year and the recent introduction of the test. Overall, the verification program appears to be functioning as described by those who originally implemented the program [7]. Data from a

1970 report [2] showed that initial failure rates for the impact verification program were 44 % (433 tests); but as machines were repaired and retested the failure rates began to decline sharply (11.5% was the lowest failure rate reported). Clearly the verification program has established and maintained a population of impact machines that can be reliably used for acceptance testing: more than 700 of the 800 machines tested annually in the program are within 1.4 J, or 5 % of the reference values, indicating that these machines differ from each other by less than 2.8 J, or 10 %.

The distributions of the data show significant numbers of machines near the limits of the pass/fail criteria, and we assume that some *good* machines failed the verification test. If the pass/fail criteria were widened to include more of these borderline machines, however, we suspect that the distribution of test results would broaden over time and a similar situation would develop near the new acceptance limits. In effect, it is the stringent pass/fail criteria adopted by E 23 many years ago that has resulted in the narrow distribution of impact test results in the program today. Frankly, any less stringent requirements would result in acceptance tests with little value, particularly when qualifying high-strength steels for severe environments. So accepting the current 1.4 J or 5 % pass/fail requirements as practical and necessary, we can strive only to decrease the probability of *good* machines failing the verification test and of *bad* machines passing the test.

The current acceptance criteria are based solely on the averages of the verification set and the pilot lot and do not take the variation of the data into account. The proposed range rule is a step in the right direction to help identify *bad* machines that would currently pass the verification test. Based on the low-energy data presented here, though, the probability of failing *good* machines with this rule because of one single outlying measurement is too high. We think it will be necessary to visually examine the specimens for jamming marks and evaluate the distribution of the five test results to more accurately identify machines with excessive variation. In addition, the proposed range rule can detect excessive variation only in the verification set and does not incorporate any of the information available on the pilot lot variation.

Further consideration of the pass/fail criteria and how they relate to the variations of the pilot lot data, our best indicator of variability, are needed. It has been proposed [5] that the candidate machine may be certified if $E_c - E_r$ is in interval (L, U) with

$$U = -L = d + 0.76 \cdot S \qquad (1)$$

where S is the pooled standard deviation of the pilot lot and d is a constant. The question is, what value should d be given. For illustration, we will use S = 0.7 J which is a typical standard deviation for the low-energy verification specimens (**Table 2**). With S = 0.7 J, U = 1.4 J, d is 0.9 J. Historically, low-energy lots have been accepted for distribution if the standard deviation of the pilot lot was 1 J or less [4], so when verification tests are conducted with specimens having a standard deviation of more than 0.7 J, say 1 J, these tests would have only to meet ±1.66 J criteria according to Eq 1. This is reasonable and would lower the probability of a *good* machine failing the verification test, but as we have already stated, to

maintain a useful verification program we have accepted a maximum difference of 1.4 J for the low-energy range. So we need to reduce the maximum allowable standard deviation for the acceptance of specimen lots to increase the probability of certifying *good* machines and leave the pass/fail criteria at their present values. This is clearly the most direct and best method by which we can improve the impact verification program.

There is always uncertainty associated with the reference energy assigned to the lot. Overall, the average energies for pilot-lot data are in good agreement with the verification test results. It is, however, very difficult to evaluate the influence of the specimens on the verification data. In practice, we mitigate the influence of the specimens on the verification test by monitoring these data and visually inspecting the five returned specimens. For example, if a test fails to meet the 1.4 J or 5 % requirement and there are no markings on the specimens to indicate that the machine is in need of repair, we will retest the machine using different specimens if we have reasons to question the specimens used in the test. In practice this approach works well in minimizing the effect of the specimens on test results, but a better estimate of the mean would help avoid this problem.

To reduce the uncertainty of the reference energy assigned to verification specimens, the sample size can be increased and/or the sampling method can be improved. We think that increasing the sample size will improve our estimate only marginally. We are, however, considering changing from a random sampling method to a systematic sampling related to the positions of specimens in the heat-treating baskets. In addition, the control specimens placed in the sampling locations would be marked to identify their bar-stock origin (the ingot location of the bar stock is known, but we do not track the individual bars from which the samples are made). This type of sampling would allow us to include two variables of our processing in a consistent manner.

Generally, most machines that perform well at the upper and lower bounds of their capacities also perform as expected at mid-range energies. But three of machines failed the mid-range test, and **Fig. 8** shows that tests at any one energy alone provide limited predictability on how the machine will perform at the other energies. For example. **Fig. 8** shows the results for a machine that had nearly perfect performance (0 % difference) at both low and super-high energies, but the machine failed the high-energy test. This is due primarily to the fact that the specimens used to test at the three energy levels interact quite differently with the machine. For this reason we believe it is necessary to test at each energy level at which the machine will be used to have any certainty of the performance of the machine at that energy level.

Conclusions

Some conclusions based on the Charpy impact verification data that were evaluated by NIST from January 1994 through December 1996 are as follows:

1. The adoption of the stringent pass/fail criteria (1.4 J or 5 %), originally proposed for verifying Charpy impact machines in 1955, has produced a large population of impact machines that are suitable for acceptance testing.

2. The range rule now being considered for E 23 can detect excessive variation in measurements. More studies are needed before the rule is implemented.
3. Impact machines of large capacity should continue to be verified by testing at three energies.

Acknowledgment

This work is a contribution of the National Institute of Standards and Technology and is not subject to copyright in the United States.

References

[1] Driscoll, D. E., "Reproducibility of Charpy Impact Test," *Symposium on Impact Testing*, *ASTM STP 176*, American Society for Testing and Materials, Philadelphia, 1955, pp. 70-74.
[2] Fahey, N. H., "The Charpy Impact Test—Its Accuracy and Factors Affecting Test Results," *Impact Testing of Metals, ASTM STP 466*, American Society for Testing and Materials, Philadelphia, 1970, pp. 76-92.
[3] ASTM E 23-88, "Standard Test Methods for Notched Bar Impact Testing of Metallic Materials," American Society for Testing and Materials, Philadelphia.
[4] ASTM E 1271-88, "Standard Practice for Qualifying Charpy Verification Specimens of Heat-Treated Steel," American Society for Testing and Materials, Philadelphia.
[5] Splett, J. D. and Wang, J. C.-M., "Proposed Changes to Charpy V-Notch Machine Certification Requirements," Pendulum Impact Machines: Procedure and Specimens for Verification, *ASTM STP 1248*, American Society for Testing and Materials, Philadelphia, 1995, pp. 182-194.
[6] John, P. W. M., *Statistical Methods in Engineering and Quality Assurance*, John Wiley & Sons, Inc., New York, 1990.

C. N. M[c]Cowan,[1] J. Pauwels,[2] G. Revise,[3] and H. Nakano[4]

International Comparison of Impact Verification Programs

Reference: McCowan, C. N., Pauwels, J., Revise, G., and Nakano, H., **"International Comparison of Impact Verification Programs,"** *Pendulum Impact Testing: A Century of Progress, ASTM STP 1380*, T. Siewert and M. P. Manahan, Sr. Eds., American Society for Testing and Materials, West Conshohocken, PA, 1999, pp. 210-219.

Abstract: A horizontal comparison is made between the four laboratories that certify Charpy impact verification specimens. The participants in this study were Japan (NRLM), France (LNE), the European Commission (IRMM), and the United States (NIST). The exercise was conducted to show how the impact verification programs, specimens, and test procedures compare with each other. Results for both 8 mm and 2 mm strikers were compared. The study showed the following: (1) The certified energies of impact verification specimens distributed by these four metrological authorities often agreed within 1 % of the average values determined in this study; (2) the variation in energy for the specimens was low, typically bracketed by a coefficient of variation of between 0.02 and 0.04; and (3) the energies measured for the tests performed with 2 and 8 mm strikers on the 4340 steel specimens were nearly equivalent, but a trend of slightly higher energy for the 2 mm tests is indicated.

Keywords: Charpy impact verification, impact testing, verification testing, verification specimens

Introduction: Charpy impact testing is often specified as an acceptance test for structural materials, and companies performing acceptance tests are typically required to verify the performance of their impact machine using certified verification specimens. To our knowledge there are only four laboratories in the world that certify and distribute reference materials for the verification of Charpy impact machines: (1) The Institute for Reference Materials and Measurements (IRMM, Belgium), (2) Laboratoire National D'Essais (LNE, France), (3) The National Institute of Standards and Technology (NIST, USA), and (4) The National Research Laboratory of Metrology (NRLM, Japan). These four laboratories supply impact verification specimens to verify the performance of an estimated 1800 impact machines annually.

[1] Materials Reliability Division, NIST, 325 Broadway, Boulder, CO, U.S.A., 80303

[2] Reference Materials Unit, EC-JRC-IRMM, Retieseweg, B-2440, Geel, Belgium

[3] Mechanical and Equipment Testing, LNE, 5 Avenue Enrico Fermi, 78197, Trappes Cedex,France

[4] Materials Measurement Section, NRLM, 1-4, Umezono 1-Chrome, Tsukuba, Ibarki, 305, Japan

This study provides the first horizontal comparison between these four laboratories. We compare both the impact verification specimens and the machines (or systems) used to certify the absorbed energies of the specimens. Our goals are to use this comparison to better understand the details of each other's verification programs, and to consider how any differences between our verification systems and specimens might affect the users of our respective programs. It is only through these types of horizontal comparisons that we can assess the equivalency of the results for impact acceptance tests made around the world. These initial results help confirm that acceptance tests performed under one system are equivalent to the others, making the verification systems and specimens transparent to the user.

Background Information on Verification Programs

The four verification programs represented in this report have similar goals and much in common, but each program is unique. To start with, there are two fundamentally different approaches used to stabilize the certification procedures for impact verification specimens. The European philosophy for stabilizing their certification procedures (EN Standards) is based on traceability to "master specimens." This system is represented here by IRMM, which currently certifies and distributes impact verification specimens for the Community Bureau of Reference (BCR). The stability of the national impact verification programs in the United States, Japan, and France is based on traceability to designated impact reference machines that are maintained by the respective metrological authorities (NIST, NRLM and LNE). This difference and details of the certification procedures discussed below (and given in **Tables 1 and 2**) make direct comparisons of our systems difficult. However, keeping these differences in mind, we can make useful comparisons of our results.

IRMM Program - In the IRMM program, "master batches" of impact verification specimens are tested in a round robin for the BCR. From the results of the round robin a certified energy for the specimens is determined, and these "master specimens" are then tested over time to track and normalize the results of the impact machine that is used to determine certified values for BCR verification specimens. The certified values for BCR impact verification specimens are determined as follows: (1) New batches of BCR impact verification specimens are tested (30 samples per batch) on a single impact machine, together with sets of "master specimens" (generally 25 or 35 specimens) of similar nominal energy; (2) if the samples have acceptable variation and energy, the difference in the average energies for the "master specimens" and the batch being evaluated is determined for the machine; (3) this difference is added or subtracted from the average energy for the batch of BCR verification specimens to determine a certified energy value[1]. These procedures meet the EN 10045 and ISO 148-3 requirements for verification specimens [2,3].

For this study, the results of the impact machine used by IRMM were reduced by 4 % to represent the BCR procedure, so these results represent the certification system, not a single impact machine. The IRMM specimens can be used to verify impact machines in accordance with EN 10045 and ISO 148-2 [4].

Table 1 - *Details on the specimens and impact machines used in the study. Under machine information, the pendulum type (C or U), the striker radius which was tested, the method of reading the absorbed energy (dial or encoder), and the maximum capacity of the machine are given.*

Laboratory	Specimen Information			Certified Values		Machines Details
	Test Temperature (°C)	Energy Level	Steel Type	Striker Radius (mm)	Value (J)	
1	20	1	4340	2	27.0	- C- type - 2 and 8 mm - Dial - 300 J
				8	26.8	
	20	2	4340	2	113.6	
				8	112.6	
	20	3	XM32	2	157.1	
2	20	1	E.40 NCD7	2	26.4	- U-type - 2 and 8 mm - Encoder - 350 J
	20	2		2	67.2	
	20	3		2	111.8	
3	-40	1	4340	8	16.5	- U- type - 2 and 8 mm - Encoder - 350 J
	-40	2		8	99.2	
	20	3	T-200	8	258.0	
4	0	1	4340	2	*	- C-type - 2 and 8 mm - Dial - 500 J
	0	2		2	*	
	0	3		2	*	

* The certified values were not available because the lot size for the specimens was too small to both certify the specimens and provide the number of specimens required for this study.

NRLM Program - In Japan, there are two C-type impact machines in the program, but a single impact machine is used to determine the certified values for the verification specimens. The second machine is used as a back-up machine and for comparison to the machine used for certifications. Because a single machine is used to determine the certified value for the verification specimens, the results for the NRLM impact machine used in this study represent those that would be attained for the Japanese certification system. The certified value for the NRLM specimens is determined by testing 25 specimens. The results are evaluated and if all statistical criteria are met, the average energy of the 25 specimens can be used as a certified value. This procedure meets the JIS B 7740 and the ISO 148-3 requirements for verification specimens (the requirements for the JIS B 7740 and ISO 148-3 standards are identical) [5].

The impact specimens certified by the NRLM program are primarily used to verify the performance of impact machines according to JIS B 7722, which is the national standard for impact testing in Japan [6]. The NRLM specimens can also be used for the ISO 148-2 machine

verification (again, the requirements for the JIS and ISO standards are identical).

NIST Program - In the United States, the certified values for impact verification specimens are determined using three impact machines (2 U-types and 1 C-type). The certified value is determined as follows: (1) 25 specimens are tested on three different impact machines, (2) the results are evaluated to determine whether the differences in the variation for the specimens and the average energies of the three machines meet established criteria, (3) if the specimen variation is acceptable and the comparison of results for the three machines are within normal bounds, the results are combined and the certified value is defined as the average value for the 75 specimens. The certified specimens meet the ASTM E 1236 Standard for Qualifying Charpy Impact Machines as Reference Machines and ISO 148-3 requirements for verification specimens. In this study, only one impact machine was used, so the NIST results represent that machine, not the certification system. However, at energies above 40 J, the results for the NIST machine are expected to closely represent the average value of the three NIST machines. For the low-energy (<40 J) specimens, the results for this single machine are often higher than the average of the three machines used to determine a certified value for NIST specimens.

Table 2. *Summary of the verification requirements for national and international standards: difference allowed for the verification result and the certified energy value (E_c).*

Designation	Requirements
ISO 148-3	±2 J for $E_c \leq 40$ J ±5% for $E_C > 40$ J
ISO 148-2	±4 J for $E_c \leq 40$ J ±10 % for $E_c > 40$ J
EN 10045	±2 J for $E_c \leq 40$ J ±7.5 % for $E_c > 40$ J
ASTM E23 ASTM 1236	±1.4 J for $E_c \leq 28$ J ±5% for $E_c > 28$ J
JIS B 7722	±4 J for $E_c \leq 40$ J ±10 % for $E_c > 40$ J
JIS B 7740	±2 J for $E_c \leq 40$ J ±5% for $E_C > 40$ J

The NIST specimens are used primarily to verify the performance of impact machines according to ASTM E 23, the Standard Test Method for Notched Bar Impact Testing of Metallic Materials. The specimens can also be used for ISO 148-2 verification tests.

LNE Program - In France, two impact machines are used to certify verification specimens (a U-type and a C-type). The machines are located at two different laboratories in France (LNE and CTA), which have an agreement from the French Community of Accreditation (COFRAC) for the certification of impact verification specimens. The procedure used to develop a certified value is in accordance with EN 10045-2 and is similar to the NIST procedure: 25 specimens are tested on each machine, and the grand average (50 specimens) is used as the certified value for the specimens if statistical evaluations show acceptable performance for the machines and the specimens. In this study only the U-type impact machine was used, so the LNE results represent that machine, which typically has a slightly higher absorbed energy than the C-type used by LNE. The LNE specimens can be used to verify impact machines according to EN 10045 and ISO 148-2.

Materials and Procedures

Each of the participating laboratories provided 100 verification specimens at three different levels of absorbed energy, referred to as levels 1, 2, and 3 in **Table 2**. The laboratories are coded using numbers (1–4). The specimens provided by each laboratory are also coded (1–4) to identify the laboratory that supplied the specimens.

Information on the impact machines used for the testing is given in **Table 2**, along with the test temperature specified for each group of specimens. The certified absorbed energies for the specimen groups, which each laboratory had previously determined for the specimens, are also given in **Table 2**. (These values were not made available to the participants until the testing was completed.)

All four laboratories used a similar type of steel to produce verification specimens for energy levels 1 and 2 (**Table 2**). The steel is a medium-carbon, low alloy, high-strength steel, designated in the U.S. as AISI/SAE 4340. Although these steels are very similar in composition, the heat treatments used by the various laboratories are assumed to vary significantly.

The types of steel used to produce verification specimens for energy level 3 differ between the various laboratories. So at this energy level, differences in impact properties due to alloy content must be considered. Basically, three types of alloys were used for energy level 3: (1) a XM32 steel, which is a Cr- Mo type alloy, (2) a T-200 steel, which is an 18 Ni-0.7 Ti maraging steel, and (3) the AISI/SAE 4340 type steel, which is also used for energy levels 1 and 2.

The impact testing was conducted using the procedures normally followed by each laboratory, and any special instructions that were provided with the specimens. From each specimen group of 25 specimens, 15 were tested using a 2 mm striker, and 10 were tested using an 8 mm striker. Some outlier data were removed during testing, but most of the statistical outliers were left in the data for our analysis.[5]

Specimen Variation

The variation in the absorbed energy for the specimen groups is given in **Table 3**. The results are presented in terms of the coefficient of variation (CV), which is the ratio of the standard deviation in the absorbed energy to the average absorbed energy of the specimen group being evaluated. This relative measure provides a convenient means of normalizing the various energy levels being considered here. A CV of about 0.04 (or less) represents a low level of the variation in absorbed energy for impact verification specimens. This CV value is related to sample size calculations (for a sample size of 5), and would generally ensure a variation small enough to conduct a verification test within the statistical considerations upon which our

---

[5] Statistical outliers are defined here as those data that were identified as outliers in box-and-whisker analysis. For the most part, data were removed during testing only due to a testing problem (specimen placement, etc.).

verification programs are designed. The ISO 148-3 requirements for verification specimens allow a CV of 0.05 or less.

The pooled CV is considered the best indicator of specimen variation, due to its inherently larger sample size. It is defined as the rms value: the square root of the sum of the squares of the CVs (for laboratories 1–4) divided by 4.

At the lowest energy level evaluated, level 1, the variation for the 8 and 2 mm tests are similar. Considering the pooled CVs, no clear effect of striker radius on the variation is apparent. Occasionally one laboratory had a high or low CV with respect to the other three laboratories. But typically the high CVs reported for the level 1 specimen groups were due to one or two statistical outliers in the data. Generally, the variation for the level 1 specimens was low, and all of the pooled CVs are 0.04 or less.

At energy level 2, the interlaboratory CVs for the various specimen groups are even more consistent than the results for level 1. Typically, three laboratories had the same CV for a specimen group, and one would differ slightly. All of the pooled CVs are 0.04 or less; many are 0.02. Again, there was no apparent effect of striker radius on the variation.

Table 3 - *The coefficient of variation is given for the four laboratories , organized by energy level, striker (tup), specimen group. The pooled CV (CVP) of the laboratories is also given. The CV is defined as the ratio of the standard deviation and the average absorbed energy for the group.*

Energy Level	TUP (mm)	Specimen Group	CVs for Labs 1–4, and the pooled CV (CVP)				
			CV1	CV2	CV3	CV4	CVP
1	2	1	0.02	0.04	0.04	0.04	0.04
1	2	2	0.02	0.03	0.02	0.05	0.03
1	2	3	0.02	0.07	0.04	0.02	0.04
1	2	4	0.02	0.02	0.02	0.03	0.02
1	8	1	0.02	0.02	0.02	0.04	0.03
1	8	2	0.03	0.04	0.03	0.06	0.04
1	8	3	0.05	0.06	0.03	0.02	0.04
1	8	4	0.02	0.03	0.04	0.01	0.03
2	2	1	0.04	0.04	0.05	0.03	0.04
2	2	2	0.02	0.02	0.02	0.01	0.02
2	2	3	0.02	0.03	0.03	0.02	0.02
2	2	4	0.02	0.02	0.03	0.02	0.02
2	8	1	0.04	0.03	0.03	0.03	0.03
2	8	2	0.02	0.01	0.02	0.02	0.02
2	8	3	0.02	0.02	0.03	0.02	0.02
2	8	4	0.02	0.03	0.03	0.03	0.03
3	2	1	0.02	0.02	0.04	0.04	0.03
3	2	2	0.03	0.04	0.03	0.04	0.04
3	2	3	0.12	0.15	0.03	0.11	0.11
3	2	4	0.03	0.03	0.03	0.04	0.03
3	8	1	0.03	0.03	0.03	0.04	0.03
3	8	2	0.04	0.04	0.03	0.02	0.03
3	8	3	0.03	0.05	0.03	0.03	0.04
3	8	4	0.05	0.05	0.04	0.05	0.05

At energy level 3, the interlaboratory CV results are slightly less consistent than the level 2 results, and differentiate the results for several specimen groups. The 2 mm tests for specimen group 3 have a significantly higher CV than the other level 3 groups (pooled CV equal to 0.11), and the 8 mm tests for specimen group 4 had interlaboratory CVs slightly higher than 0.04 (pooled CV equal to 0.05). Both of these specimen groups, however, had acceptable CVs for

tests made using the striker radius for which they were developed: The CVs for the 8 mm tests on the group 3 specimens and the 2 mm tests on the group 4 specimens were 0.04 and 0.03 respectively. So, for the more ductile high-energy specimens, striker radius apparently had an effect on the variation in absorbed energy.

In general, the interlaboratory CV for the various specimens groups were in good agreement and appear to reflect differences in the specimens, not the impact machines. However, laboratory 1 had the lowest variation most often, and this likely indicates some influence of the impact machine on the variation assigned to the specimens.

Absorbed Energy

The absorbed energy results for the specimen groups are shown in Figures 1- 6 and the data are given in the Appendix. The plots show the average absorbed energy for each specimen group, with error bars representing ± 1 standard deviation. The results for each energy level are shown separately for the 2 and 8 mm tests. For comparison purposes, a fifth average and error bar was added to the data for each specimen group. This bar shows the average absorbed energy of the four laboratories for each group (the grand average) and the ASTM E23 and EN 10045 limits for verifying the performance of reference quality impact machines (see Table 1). These error bars offer convenient scales for comparisons of these data, but the limits have no real significance for the horizontal comparisons we make here. The data for each laboratory, within each specimen group, are plotted from left to right as laboratory 1, 2, 3, and 4.

Energy Level 1- As shown in Figures 1 and 2, laboratory 4 always had the lowest absorbed energy for the specimen groups tested at this energy level. The highest values were always from laboratory 2 or 3, and laboratory 1 always had the intermediate values.

The relative relationships between these results were found consistent for both the 2 and 8 mm test results, but the magnitudes of the differences between the results were dependent on the particular specimen group that was tested. The apparent differences in the results for the laboratories were greatest for the tests made with specimen groups 1 and 3, slightly smaller for specimen group 2, and smallest for specimen group 4 (±0.8 J).

The average results for the tests were all within ± 2 J of the grand average determined for the energy level 1 specimen groups, and most were within ± 1.4 J of the grand averages. So these results are within the limits deemed acceptable for reference-quality impact machines. The range in the results, however, pushes these limits. This is partly due to differences in the designs of the C and U-type impact machines represented here. The impact machine used by laboratory 4, for example, may be the most rigid machine in this study (large capacity C-type machine) and the lower values produced by this machine may be very good values (less energy absorbed by the machine in the test).

It must also be kept in mind here that the average values for two of the laboratories would likely be lower if the C-type machines that are also used in their verification programs were included in their averages. This would decrease the range in the results.

The 2 and 8 mm test results were equivalent, considering the standard deviations associated with them. But there is a trend for the 2 mm results to have slightly higher absorbed energies. The average differences between the 8 and 2 mm results for each specimen group were as follows: (G1) -0.01 J, (G2) +0.13 J, (G3) +0.44 J, and (G4) +0.06 J. The pooled standard deviations associated with the average energies for these specimen groups were in the range of 0.5 to 1.0 J.

Energy Level 2 - The results of the laboratories for the level 2 specimen groups (Figures 3 and 4) are well within ± 5 % of the grand averages determined for the groups (none exceed ± 3 %). The results do not show significant differences between the laboratories, and the effect of striker radius is small. In general the 2 mm test results have slightly higher energy than the 8 mm results: the average differences between the 2 and 8 mm tests for the laboratories were 1.6, 1.1, 1.1, and 1.5 J for the specimen groups 1-4 respectively. In most cases, the differences are less than the standard deviations associated with the average energies of the specimen groups being compared.

Energy Level 3 - The tests at energy level 3, Figures 5 and 6, show a significant difference in the results for the 2 and 8 mm striker radius. The differences between the 2 and 8 mm test results are apparently related to the type of steel used for the specimens as much as they are to the energy level of the specimens. For the specimens made from 4340-type steel (groups 2 and 4), the differences in the 2 and 8 mm tests were the smallest. The group 2 results were mixed with respect to the influence of the striker radius and the differences (average difference of −0.7 J) were not statistically significant. The results for the group 4 specimens consistently show a higher energy for the 8 mm tests, and the differences (average difference of −4.8 J) are on the order of one standard deviation. For the specimens made from the XM32 steel (group 1), the 2 mm test results were consistently higher than those for the 8 mm tests. The average difference for the group 1 specimens was +10.0 J, which is statistically significant (standard deviations of the average energies for group 1 were around 5 J). The tests with the specimens made from the maraging steel (group 3) had the largest difference in 2 and 8 mm results (average difference of −44.7 J), and this shows the results for the 8 mm tests to be higher than those for the 2 mm tests. So the results for two steels (XM32 and maraging) show significant differences for 2 and 8 mm test results, but the effect of the striker radius was quite different (opposite).

The results also show that the group 3 maraging-steel specimens are better verification specimens for the 8 mm test, for which they were developed. The variation in the absorbed energies for the 2 mm tests were much higher than those for the 8 mm tests, and the differences in the interlaboratory average energies were also greater for the 2 mm tests. So it appears that the group 3 maraging-steel specimens perform adequately as 8 mm verification specimens, but are a poor choice for 2 mm verification tests. The group 1, 2, and 4 specimens all performed well as verification specimens for level 3, for both 2 and 8 mm tests (all within the ±7.5 % limit, and most within the ±5% limits).

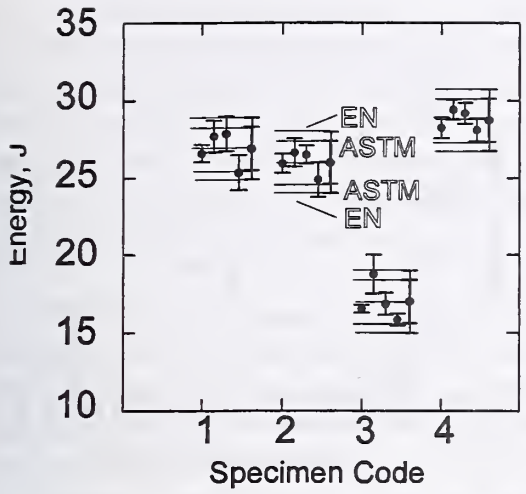


Figure 1. Energy level 1, 2 mm striker.

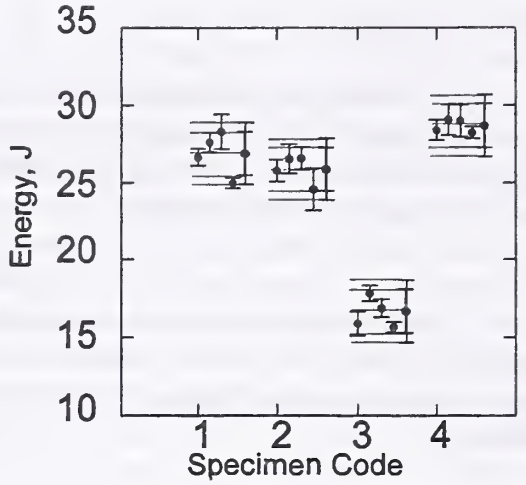


Figure 2. Energy level 1, 8 mm striker.

Figure 3. Energy level 2, 2 mm striker.

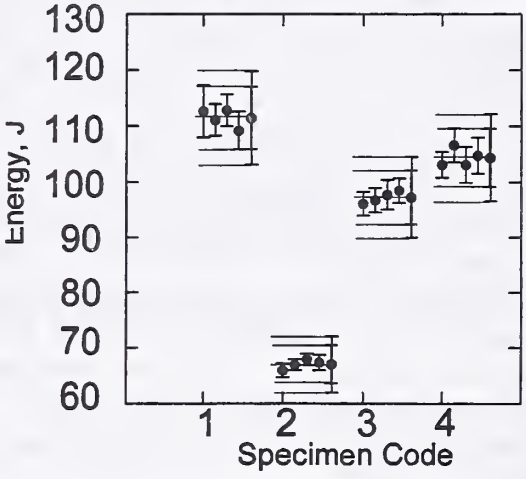


Figure 4. Energy level 2, 8 mm striker.

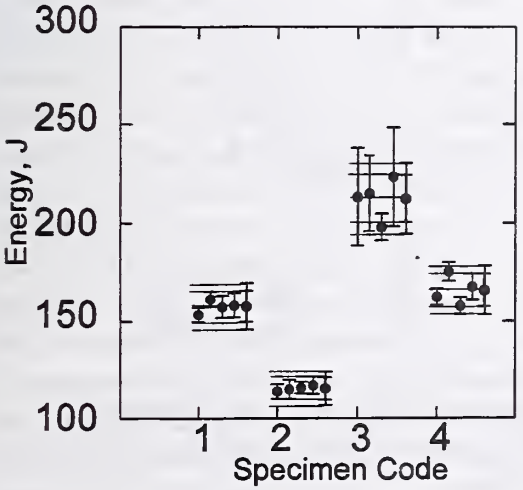


Figure 5. Energy level 3, 2 mm striker.

Figure 6. Energy level 3, 8 mm striker.

Lateral Expansion

The lateral expansion for the specimens, **Figure 7**, increased with increasing absorbed energies, as expected. The correlation is linear up to about 150 J. The 2 mm tests resulted in slightly increased lateral expansion (and absorbed energy) for most of the specimen groups. The exception was specimen group 3 (level 3), where the 8 mm tests had higher lateral expansion. This result is consistent with the higher absorbed energies for the 8 mm tests on these specimens.

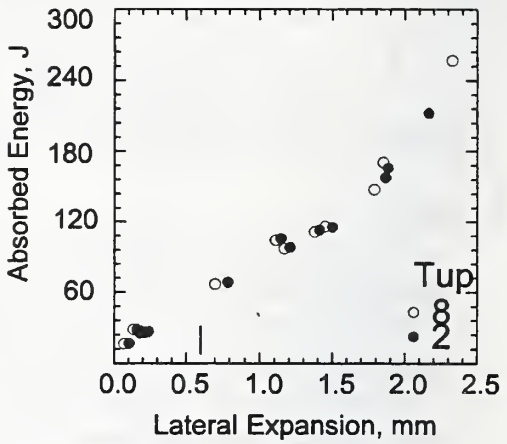


Figure 7- *The grand averages for the lateral expansion and absorbed energy for each specimen group are plotted here (averages of the 4 laboratories for each group). Results for the 2 and 8 mm tests are shown.*

Hardness

As shown in **Figure 8**, the absorbed energy decreases with increasing hardness. The relationship is approximately linear, if only the type 4340 steel specimens are considered. The XM32 steel fits the trend, but the T200 maraging alloy does not. This result is not surprising, but the plot does provide a good example of the relationship between hardness and absorbed energy for the impact verification specimens. The variation in hardness measured for the specimen groups was generally low. Reviewing these data it appears that a standard deviation in hardness of between 0.1 and 0.2 HRC is routinely attained for samples at all energy levels. This corresponds to a range in the hardness of 1 HRC for a specimen group (25 specimens). The specimens having the most consistent hardness, however, did not necessarily have the lowest variation in absorbed energy.

There was a consistent offset in the hardness values reported by the various laboratories. Since hardness measurements were not the focus of this study (not mandatory), this result is not of direct concern here. However, it is of general interest that the average hardnesses for the specimen groups measured by the four laboratories were almost never within 1 HRC, and sometimes differed by more than 2 HRC.

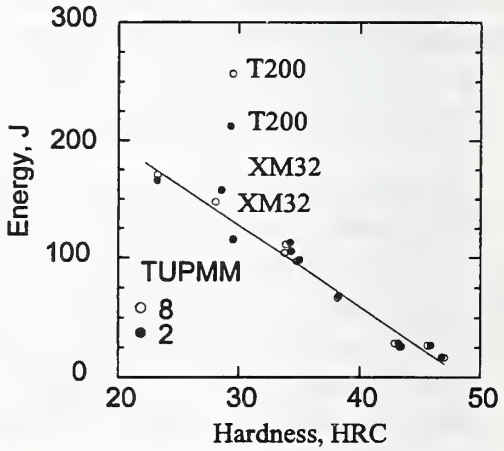


Figure 8 - *Hardness and absorbed energy of the steels used for vertification testing.*

Discussion

The specimens used for verifying Charpy impact machines around the world have similar variation in absorbed energy, which is the principal

factor used to evaluate the quality of these specimens. The coefficient of variation for the specimens from all laboratories and energy levels typically ranged from 0.02 to 0.04. This result is interpreted as a benchmark for the quality of impact verification specimens for the 1990s. Machine/specimen interactions (and other variables) affect the variation assigned to a particular specimen group by a single machine, but in general the differences in specimen variation were resolvable and appear to reflect the homogeneity in the impact properties, rather than effects of machines, operators, and sampling.

The relative outcome of the absorbed energy results for the laboratories was sometimes dependent on the specimen group that was tested. This was particularly true for the tests made using the lowest energy specimens (group 1), where the changes in interlaboratory results due to the specimens altered the general interpretation of the results. For example, considering only the tests made with specimen groups 1 and 2, one might conclude that the results from laboratory 4 were different and lower than those of the other laboratories, but considering only the results for specimen group 3, the results from laboratory 2 appear to be different and higher than those of the other laboratories. If only the results for the group 4 specimen were considered, we would conclude that the laboratories all produce similar results for low-energy impact tests. These specimen effects are secondary in many cases, but raise questions concerning what these apparent differences in the results show us about the performance of the impact machines tested. Because these specimen effects are not well understood for the most part and are not clarified by these results, we refrain from speculating here. The data do, however, provide a useful side-by-side comparison of these differences for our verification specimens and will hopefully lead to discussions and testing that improves our understanding of specimen/machine interactions.

The summary plot of the 2 and 8 mm tests results in **Figure 9** shows that the effect of striker radius is small (statistically insignificant) for the specimens made from 4340 steel. The small differences in energy for the 2 strikers are not consistent over the 10 to 150 J range: they show the tests results for the 2 mm striker to be slightly higher for the specimens at energy levels 1 and 2 (average difference of +0.16 and +1.33 J respectively), but lower at energy level 3 (average difference of −2.75 J). The lateral expansion results indicate a more consistent trend of increased ductility (higher toughness) for the 2 mm tests over the range of 10 to 150 J for the 4340 specimens. This finding is consistent with several previous studies that evaluated these effects for 4340 steels [7−11]. But, for certifications even small differences in the test results must be considered. So, we conclude that separate certifications for 2 and 8 mm tests are needed for verification specimens made from 4340 steel.

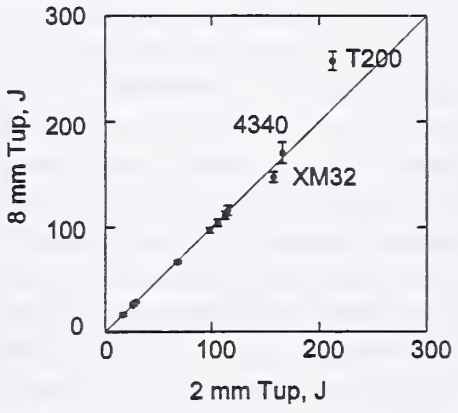


Figure 9 - *The 2 and 8 mm striker results (grand averages) are shown. The grand average of the 2 and 8 mm strikers include about 60 and 40 tests respectively. The error bars represent ±1 standard deviation for the 8 mm averages.*

The results for the XM32 and T200 steels were strongly affected by the striker radius used for the test. For the very ductile T200 material, this greater sensitivity has been explained by the fact that the larger radius 8 mm striker (and striker design) requires more bending of the specimen during fracture and this requires more energy [7,8,10]. These same mechanical differences may also explain the better performance of the T200 specimens with the 8 mm striker: the increased bending results in more complete tearing during fracture, which results in a more consistent fracture energy for the specimens. The higher energy for the 2 mm tests on the XM32 material, however, has not yet been convincingly explained. Studies with other materials, showing similar effects of striker radius, have speculated that this increased energy for 2 mm tests may be due to deeper penetration of the smaller striker combined with mechanical properties specific to the material (strain hardening, strength, hardness, etc.) [10]. Here, the lateral expansion for the 2 mm tests on the XM32 samples was lower than those for the 8 mm tests (and similar to the 2 and 8 mm results for the 4340 steel), so considering the ductility associated with the fracture does not provide an explanation.

Overall, the comparisons of our results are reassuring. Even though the different specimen groups clearly had an effect on some test results, the trends for the test results were generally independent of the specimens tested. The average absorbed energies determined by the four laboratories for the various specimen groups, and the variation in absorbed energies, are in good agreement. The laboratories differed mostly on results for verification specimens of the lowest and highest energy (energy levels 1and 3). But the results never differed from the average by more than 2 J (or 7.5 %). In most cases the ASTM 1236 requirement for qualifying reference machines of 1.4 J (or 5 %) was met, and if the verification systems could have been more directly comparable (by including C-type machine results in the averages for laboratories 2 and 3), the range in absorbed energies between the verification laboratories would likely decrease.

Our laboratories have long recognized that differences in the designs of impact machines have an effect on impact test results, and our verification systems include this factor (with the exception of Japan) when determining a certified energy for impact verification specimens. Japan uses a single C-type machine for certifying verification specimens, but the vast majority of the companies verifying impact machines in Japan also have a C-type machine (so Japan does not necessarily have to consider this variable). In Europe and the United States, however, many different types of impact machines are used by our industries, so we try to include the differences due to machine design in our certification procedures. As explained in the background section, the certified energy for ASTM verification specimens is determined by using the average energies from three different impact machines (2 U-types from different manufactures and 1 C-type). The French system uses one C-type and one U-type machine to determine certified values for verification specimens. The verification system used by the European Commission (BCR) uses the results from round-robin testing on master specimens (which includes many different impact machine designs) to adjust the result of reference machines used to assign certified values to verification specimens. This practical approach of averaging the machine design effects has worked reasonably well, but the limits specified for verification tests are quite stringent in some cases, and the machine design effects can be significant, particularly when testing low-energy specimens (as shown by our results). However, considering the ISO 148-2 requirements for verifying the performance of impact machines (which allow a ±4 J error for test at energies of

40 J or less, and a 10 % error for tests at energies over 40 J), the agreement found for our systems appears to be adequate for purposes of international commerce. As impact machines of newer design and larger capacity are produced we will have to reevaluate this approach (and our respective limits on verification of test results).

As a final point for comparison, we consider how well the certified energies for the specimen groups tested in this study agree with the average energies determined by our laboratories. As shown in **Figure 10**, the agreement is good. Again, it was difficult to make a true comparison here, because the certified values for the Japanese specimens were not available. However, the results for the Japanese impact machine in this study were substituted for their certified values, and these values are expected to closely estimate the certified values. For the 17 cases compared here, the average energy determined by the laboratories in 12 cases differed from the certified value by less than 1.0 %, 4 differed by 1.0 to 2.2 %, and 1 differed by 3.2 %.

This is a comforting result, which indicates that the certified values being determined by the laboratories independently, are virtually identical to the grand averages determined by the four laboratories in this study. So, we believe that industries using our specimens to verify the performance of Charpy impact machines to the requirements of ISO 148-2 (which is the machine verification document we all have in common) can be assessed in a fair and meaningful manner.

Conclusions

1) The certified energies assigned to verification specimens by the laboratories which supplied them for this study agree closely with the average energies determined by the four participating laboratories.

2) The impact verification specimens distributed by all of the laboratories are similar in quality, based on the similar variations in energy found for the specimens tested in this study. The variation in the energy of impact verification specimens is currently controlled within a CV range of 0.02 to 0.04.

3) The specimens used to verify the performance of a Charpy impact machine can affect the apparent performance of the machine.

4) The absorbed energies measured with 2 and 8 mm strikers are very similar for 4340 steel, but can differ significantly for other materials used to produce verification specimens.

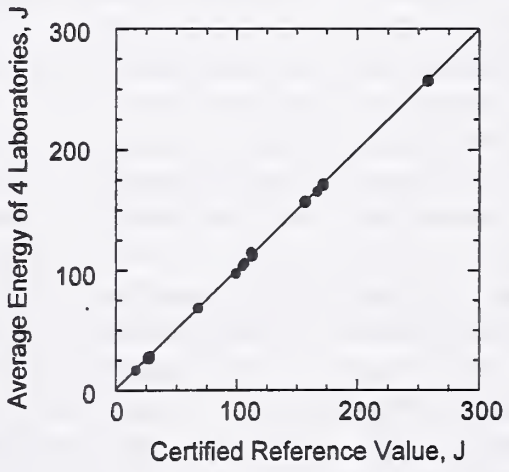


Figure 10. Average energy of the four laboratories versus the certified energy of the laboratory that provided the specimens.

Acknowledgments

C. N. McCowan would like to thank D. P. Vigliotti and T. A. Siewert from the NIST Materials Reliability Division for their efforts in measuring and testing the specimens, and in the preparation of this manuscript. J. Pauwels wants to thank Mr. Gyppaz, responsible of the Quality Department of Cogne Acciai Speciali srl, Aosta (IT) for carrying out the IRMM/BCR measurements and Dr. C. Ingelbrecht, responsible for physical reference materials within the IRMM Reference Materials unit for his valuable assistance in the evaluation of the results.

References

[1] Pauwels, J., Gyppaz, D., Varma, R., and Ingelbrecht, C., European Certification of Charpy Specimens: Reasoning and Observations, Pendulum Impact Testing: A Century of Progress, ASTM STP 1380, T. A. Siewert and M. P. Manahan, Sr., Eds., American Society for Testing Materials, West Conshohocken, PA, 1999.

[2] EN 10045-2; Metallic Materials - Charpy Impact Test - Part 2: Verification of Pendulum Impact Testing Machines.

[3] ISO 148-3, Metallic Materials - Charpy Pendulum Impact Test - Part 3: Preparation and Characterization of Charpy V Reference Test Pieces for Verification of Test Machines, International Organization for Standards, Geneva, Switzerland.

[4] ISO 148-2, Metallic Materials - Charpy Pendulum Impact Test - Part 2: Verification of Test Machines, International Organization for Standards, Geneva, Switzerland.

[5] JIS B 7740, Charpy Pendulum Impact Test - Reference Test Pieces for Machines, Japanese Institute of Standards, Japan.

[6] JIS B 7722, Charpy Pendulum Impact Test - Verification of testing machines, Japanese Institute of Standards, Japan.

[7] Ruth, E. A., Striker Geometry and its Effect on Absorbed Energy, Pendulum Impact Machines: Procedures and Specimens for Verification, ASTM STP 1248, T. A. Siewert and A. K. Schmider, Eds., American Society for Testing Materials, Philadelphia, 1995, p 101-110.

[8] Nanstad, R. K. and Sokolov, M. A., Charpy Impact Test Results on Five Materials and NIST Verification Specimens Using Instrumented 2 and 8 mm Strikers, Pendulum Impact Machines: Procedures and Specimens for Verification, ASTM STP 1248, T. A. Siewert and A. K. Schmider, Eds., American Society for Testing Materials, Philadelphia, 1995, p 111-139.

[9] Revise, G., Influence of Dimensional Parameter of an Impact Test Machine on the Results of a Test, Charpy Impact Test: Factors and Variables, ASTM STP 1072, J. M. Holt, Eds., American Society for Testing Materials, Philadelphia, 1995, pp. 35- 53.

[10] Siewert, T. A. and Vigliotti, D. P., The effect of Charpy V-Notch Striker Radii on the Absorbed Energy, Pendulum Impact Machines: Procedures and Specimens for Verification, ASTM STP 1248, T. A. Siewert and A. K. Schmider, Eds., American Society for Testing Materials, Philadelphia, 1995, pp. 140-152.

[11] Yanaka, M., et. al., Effects of the Striking Edge Radius and Asymmetrical Strikes on Charpy Impact Test Results, Pendulum Impact Machines: Procedures and Specimens for Verification, ASTM STP 1248, T. A. Siewert and A. K. Schmider, Eds., American Society for Testing Materials, Philadelphia, 1995, pp. 153-170.

Appendix

Data Summary

LEVEL	SPECIMEN	LAB	TUP	EMEAN (J)	ESTD (J)	LEMEAN (mm)	LESTD (mm)	HMEAN (HRC)	HSTD (HRC)
1	1	1	2	26.60	0.56	0.16	0.02	46.32	0.18
1	1	2	2	27.71	1.03	0.32	0.04	45.92	0.20
1	1	3	2	27.88	1.13	0.22	0.02	45.69	0.14
1	1	4	2	25.35	1.14	0.27	0.04	45.78	0.10
1	1	1	8	26.66	0.55	0.12	0.03	46.35	0.13
1	1	2	8	27.62	0.63	0.22	0.02	45.00	0.00
1	1	3	8	28.31	1.12	0.19	0.03	45.69	0.10
1	1	4	8	24.99	0.34	0.22	0.02	45.90	0.11
1	2	1	2	25.95	0.61	0.16	0.01	44.03	0.09
1	2	2	2	26.65	0.90	0.26	0.02	42.82	0.39
1	2	3	2	26.51	0.58	0.22	0.03	43.64	0.18
1	2	4	2	24.89	1.13	0.20	0.03	43.58	0.15
1	2	1	8	25.79	0.71	0.12	0.02	44.20	0.20
1	2	2	8	26.55	0.94	0.22	0.02	42.15	0.34
1	2	3	8	26.58	0.72	0.20	0.02	43.68	0.23
1	2	4	8	24.58	1.37			43.48	0.17
1	3	1	2	16.52	0.25	0.04	0.01	47.75	0.21
1	3	2	2	18.75	1.26	0.19	0.05	46.33	0.99
1	3	3	2	16.83	0.72	0.08	0.02	46.57	0.36
1	3	4	2	15.83	0.38	0.11	0.01	47.10	0.34
1	3	1	8	15.87	0.78	0.02	0.01	47.70	0.20
1	3	2	8	17.81	0.50	0.10	0.02	46.60	1.70
1	3	3	8	16.85	0.58	0.07	0.02	46.54	0.25
1	3	4	8	15.64	0.33	0.09	0.01	47.09	0.33
1	4	1	2	28.25	0.67	0.11	0.01	43.90	0.16
1	4	2	2	29.43	0.64	0.21	0.03	43.00	0.00
1	4	3	2	29.19	0.68	0.13	0.01	42.91	0.10
1	4	4	2	28.11	0.74	0.18	0.02	43.29	0.12
1	4	1	8	28.41	0.66	0.09	0.01	43.85	0.17
1	4	2	8	29.08	0.98	0.18	0.02	41.89	0.22
1	4	3	8	29.00	1.03	0.11	0.02	42.92	0.10
1	4	4	8	28.25	0.37	0.17	0.02	43.22	0.15
2	1	1	2	112.13	4.59	1.31	0.05	34.88	0.16
2	1	2	2	113.88	4.00	1.47	0.04	33.88	0.30
2	1	3	2	114.92	5.69	1.46	0.06	34.08	0.21
2	1	4	2	111.23	3.50	1.41	0.05	34.03	0.18
2	1	1	8	112.66	4.66	1.28	0.06	34.92	0.18
2	1	2	8	111.12	2.86	1.39	0.04	32.65	0.67

LEVEL	SPECIMEN	LAB	TUP	EMEAN (J)	ESTD (J)	LEMEAN (mm)	LESTD (mm)	HMEAN (HRC)	HSTD (HRC)
2	1	3	8	112.84	2.84	1.41	0.05	33.98	0.17
2	1	4	8	109.19	3.41	1.41	0.05	33.89	0.25
2	2	1	2	66.84	1.51	0.73	0.04	38.97	0.09
2	2	2	2	70.23	1.68	0.90	0.09	37.63	0.35
2	2	3	2	68.49	1.08	0.74	0.06	38.42	0.17
2	2	4	2	67.96	1.00	0.78	0.04	38.15	0.11
2	2	1	8	65.91	1.21	0.65	0.03	38.98	0.14
2	2	2	8	66.92	0.96	0.73	0.04	37.10	0.21
2	2	3	8	67.89	1.02	0.73	0.04	38.45	0.10
2	2	4	8	67.32	1.36			38.03	0.15
2	3	1	2	96.28	1.78	1.10	0.04	36.17	0.24
2	3	2	2	99.31	2.50	1.26	0.05	33.97	0.61
2	3	3	2	98.87	2.64	1.21	0.04	35.11	0.20
2	3	4	2	99.15	2.05	1.25	0.04	34.92	0.30
2	3	1	8	96.11	2.14	1.11	0.04	36.03	0.34
2	3	2	8	96.72	2.20	1.19	0.06	33.65	0.53
2	3	3	8	97.74	2.65	1.14	0.04	34.95	0.37
2	3	4	8	98.47	2.21	1.22	0.03	35.03	0.18
2	4	1	2	104.39	2.50	1.07	0.04	34.98	0.06
2	4	2	2	107.72	2.04	1.20	0.04	34.00	0.33
2	4	3	2	104.94	3.11	1.15	0.04	33.88	0.15
2	4	4	2	106.34	2.50	1.15	0.05	34.24	0.14
2	4	1	8	103.10	2.32	1.04	0.03	34.95	0.16
2	4	2	8	106.59	3.02	1.16	0.05	31.95	0.16
2	4	3	8	103.09	3.25	1.08	0.05	33.95	0.09
2	4	4	8	104.77	3.22	1.15	0.04	34.29	0.13
3	1	1	2	153.61	3.74	1.79	0.11	29.47	0.09
3	1	2	2	161.46	3.92	1.94	0.12	28.67	0.36
3	1	3	2	157.56	5.59	1.92	0.08	28.29	0.23
3	1	4	2	158.27	6.19	1.83	0.06	28.18	0.18
3	1	1	8	147.74	3.97	1.72	0.08	29.10	0.17
3	1	2	8	149.98	4.38	1.85	0.09	26.85	0.34
3	1	3	8	146.08	4.26	1.82	0.07	40.97	40.35
3	1	4	8	146.92	6.08	1.77	0.26	28.43	0.18
3	2	1	2	113.93	3.81	1.41	0.06	30.78	0.16
3	2	2	2	114.98	4.99	1.55	0.05	28.03	1.94
3	2	3	2	115.82	2.99	1.51	0.05	29.97	0.26
3	2	4	2	116.96	4.45	1.53	0.05	29.82	0.15
3	2	1	8	115.58	4.52	1.37	0.05	30.85	0.17
3	2	2	8	120.16	5.01	1.54	0.09	27.45	0.86
3	2	3	8	115.07	3.53	1.44	0.04	30.13	0.40
3	2	4	8	113.55	2.25	*		29.88	0.16
3	3	1	2	213.29	24.61	2.06	0.09	30.33	0.18

LEVEL	SPECIMEN	LAB	TUP	EMEAN (J)	ESTD (J)	LEMEAN (mm)	LESTD (mm)	HMEAN (HRC)	HSTD (HRC)
3	3	2	2	208.17	18.80	*		29.00	0.00
3	3	3	2	198.04	6.82	2.17	0.07	29.25	0.30
3	3	4	2	223.34	25.00	2.21	0.16	29.12	0.18
3	3	1	8	255.46	7.56	2.25	0.12	30.30	0.20
3	3	2	8	259.91	13.14	*		29.11	0.22
3	3	3	8	254.99	7.13	2.27	0.04	29.28	0.16
3	3	4	8	257.93	7.53	2.47	0.04	29.33	0.16
3	4	1	2	162.60	4.15	1.81	0.07	24.55	0.33
3	4	2	2	175.47	4.81	1.94	0.04	22.10	0.47
3	4	3	2	157.88	4.35	1.84	0.04	23.27	0.14
3	4	4	2	167.61	6.69	1.97	0.17	23.30	0.17
3	4	1	8	170.24	9.23	1.79	0.06	24.70	0.26
3	4	2	8	178.40	8.80	1.94	0.06	22.03	0.61
3	4	3	8	162.31	6.37	1.79	0.03	23.22	0.15
3	4	4	8	171.69	8.20	1.89	0.04	23.28	0.16

* Not fractured completely

C. Ming Wang[1] and Jolene D. Splett[1]

CONSENSUS VALUES AND REFERENCE VALUES ILLUSTRATED BY THE CHARPY MACHINE CERTIFICATION PROGRAM

REFERENCE: Wang, C. M., and Splett, J. D., "Consensus Values and Reference Values Illustrated by the Charpy Machine Certification Program," *Journal of Testing and Evaluation*, JTEVA, Vol. 25, No. 3, May, 1997, pp. 308–314.

ABSTRACT: We present an overview of consensus values and reference values for various situations and apply the computational methods to data from the Charpy machine certification program administered by the National Institute of Standards and Technology. Candidate Charpy machines are certified by comparing a reference value to the average of a set of verification specimens measured by the candidate machine. Currently, the reference value is the average of three averages observed for measurements taken using three Charpy reference machines. However, the simple average, which does not account for differences among reference machines, may not be the optimal method of computing the reference value for every situation. We describe four different methods based on the simple average, the average weighted by sample size, the weighted average that takes into account machine differences, and the weighted average when machine biases are known, for computing the consensus value and its uncertainty. We also present two methods for computing reference values using weighted and unweighted averages. The techniques and recommendations described in the paper are applicable to consensus value and reference value problems in general.

KEYWORDS: consensus value, reference value, weighted averages, variance components, notched–bar testing, reference specimens, Charpy machine certification program, pendulum impact machines, impact testing

---

[1]Mathematical Statistician, Statistical Engineering Division, Information Technology Laboratory, National Institute of Standards and Technology, Boulder, CO 80303.

Sometimes it is necessary to combine data from independent testing methods to determine a single quantity that represents all the data collected. Such is the case in the Charpy Machine Certification Program administered by the National Institute of Standards and Technology (NIST). The implementation of the program depends on the computation of a reference value which represents data collected using three different Charpy reference machines. The reference value quantifies the breaking strength of a sample of Charpy specimens when no particular reference machine is known to give the true value.

Basically, the Charpy Machine Certification Program works as follows. NIST obtains a pilot lot of about 100 Charpy specimens from a supplier and measures the impact toughness of the specimens using one of three Charpy reference machines so that roughly 33 specimens are tested on each machine. Impact toughness is measured as the energy absorbed by the specimen (in joules, J) during the test. If a pilot lot is acceptable, according to contract specifications set forth by NIST, the remainder of the lot (about 1100 specimens) is machined and delivered to NIST. NIST then sells the lot of 1100 *reference* specimens in sets of five to companies who wish to have their Charpy machines certified. The NIST procedure for qualifying Charpy specimens as reference specimens is similar to ASTM Standard Practice E 1271-88 [1].

To certify a candidate machine, five reference specimens obtained from NIST are broken using the machine under test and the results are sent back to NIST for analysis along with the broken specimens. A candidate machine is certified, according to limits specified in ASTM Standard Practice E 23-88 [2], if the average energy absorbed by the five reference specimens broken on the machine under test is within 1.4 J, or 5% of the reference value, whichever is greater. The reference value for a single lot is computed as the average of the individual reference machine averages based on the associated pilot lot data. NIST currently certifies Charpy machines for "Low" and "High" energy ranges; the reference value for a typical Low energy lot is between 15.4 and 21.0 J, while High energy reference values are usually between 92.4 and 109.2 J.

The certification process can be improved through careful study of the statistical procedures currently in use. For example, the certification limits described in the previous paragraph (the average absorbed energy of the candidate machine is within 1.4 J or 5% of the reference value, whichever is greater) are arbitrary, and alternate limits that take into account pilot lot variability have been proposed [3]. However, the alternate certification limits are applicable only if we can determine an appropriate reference value and its uncertainty.

The purpose of this paper is to describe different ways of combining data from

independent testing methods for various situations. We first present four methods for computing a consensus value and its uncertainty when there are significant differences among the reference machine averages, standard deviations, and sample sizes. The four consensus values are the simple average, the weighted average based on sample size, the weighted average that takes into account machine differences, and the weighted average based on known machine biases. Next, we discuss two reference values based on weighted and unweighted averages of individual machine averages. Reference value and consensus value problems are similar; however, there is an important distinction between them which will be discussed in detail in the Model section of this paper. The consensus value and reference value computations are applied to data from the Charpy machine certification program, and guidelines for use are given.

Although we have used the Charpy Machine Certification Program to illustrate the methods discussed in this paper, consensus value and reference value computation is common in many fields and the techniques demonstrated here will apply to similar problems in general.

The Model

The concept of combining data from independent testing methods is intuitively appealing; however, it is important to formally define the problem in a statistical framework if we are to examine various procedures for computing a consensus or reference value. The basic statistical model we will use to describe pilot lot measurements from Charpy reference machines can be written as

$$Y_{ij} = \mu + b_i + \epsilon_{ij} \tag{1}$$

where Y_{ij} represents the jth measurement observed for the ith machine. In addition, $i = 1, 2, \cdots, m$, where m is the number of reference machines on which measurements are taken (3 in our case), and $j = 1, 2, \cdots, n_i$, where n_i is the number of specimens broken by the ith machine. For example, a typical pilot lot of 100 specimens might have the following number of specimens broken on $m = 3$ reference machines.

Machine (i)	Number of Specimens (n_i)
1	33
2	34
3	33

The quantity μ is the true average energy absorption inherent in the whole lot of specimens, b_i is the amount by which the true response of the ith machine deviates

from the lot average μ, that is, the effect or bias of the ith machine, and ϵ_{ij} is the error associated with a single value. In other words, each observation can be thought of as an overall average plus a machine bias plus random error. We will assume that the ϵ_{ij} are all independent with average 0 and variance σ_i^2. Sometimes, it is assumed that all σ_i^2 are equal, but we will allow the reference machines to have different error variances.

It is critical to understand the various model assumptions and their impact on the analysis. Two basic assumptions can be made about the reference machines used in the study: 1) the machines are a random sample from a population of Charpy machines (random–bias assumption), and 2) the machines are not a sample at all but comprise a "fixed" (finite) population of interest (fixed–bias assumption). These two assumptions define the scope of the study. For example, if we assume our reference machines are a random sample from a population of Charpy machines, then the value of μ we are estimating pertains to the *entire population of Charpy machines.* In contrast, if our population of reference machines is fixed, the value of μ in eq. (1) is the mean of the *three reference machines.* Thus the scope of the study, estimating the mean of the population versus estimating the mean of a fixed set of machines determines whether the estimated mean is a consensus value or a reference value. If estimating the mean of the population is required, then our estimator is called the consensus value, while reference values are computed when the reference machines are fixed.

Most authors [4, 5, 6, 7, 8] study consensus values based on the random–bias assumption, so the b_i are normally distributed with average 0 and variance σ_b^2. The quantity σ_b^2 is a measure of the variation among the population of Charpy machines. A third assumption, that the reference machines are a *nonrandom sample* from the population of Charpy machines, is a combination of the two previous assumptions since the reference machines are fixed but the goal of the study is to estimate the mean for the population of Charpy machines. Eberhardt, Reeve, and Spiegelman [9] derive a consensus value based on the third assumption that the biases b_i are *fixed* and their bounds are known.

One feature of the random–bias assumption is that when the number of machines m is very small (say, 2 or 3), the number of degrees of freedom (usually $m-1$) associated with the estimate of σ_b^2 is also very small, resulting in an unstable estimate of σ_b^2. Since an estimate of σ_b^2 is used in the calculation of the standard error of the estimated μ, or in the estimation of μ itself, the random–bias approach is most useful when m is relatively large. However, the random–bias assumption can be inappropriate even when m is large if the m machines are not representative of the population of Charpy machines. Although the fixed–bias assumption is not affected by small values of m, it

may not be practical in some situations to completely specify the bias bounds required to compute the consensus value.

In the Charpy machine certification program, the three reference machines used to evaluate a pilot lot are *not* randomly selected from the population of Charpy machines, so the fixed–bias assumption applies. In addition, our estimate of μ in eq. (1), defined as the true average energy absorption inherent in the whole lot of specimens, applies to the three reference machines rather than the entire population of Charpy machines as in the random–bias case. In other words, the purpose of the Charpy program is to compare candidate machines to the three reference machines, so we are not interested in estimating the mean of the entire population of Charpy machines; we are estimating the mean of the three reference machines so that candidate machines can be adjusted to match the values produced by the reference machines. Thus our estimate of μ in the Charpy program is a reference value.

Although estimating μ in the Charpy program is clearly a reference value problem, we will compute the various consensus values using the Charpy data for illustration.

Consensus Values and Weighted Averages

There are many possible estimators of the true absorbed energy μ based on the model in eq. (1). In this section we will describe four different consensus values (estimators of μ) when data has been collected using different Charpy reference machines. Specifically, a discussion of three random–bias estimators will be followed by the presentation of the fixed–bias estimator of Eberhardt et al.

We will use μ to denote the true, unknown value of the population parameter (the true average absorbed energy of the lot), while $\hat{\mu}$ represents our parameter *estimate*, or consensus value, calculated from observed data. All proposed estimators of μ have the form of a weighted average of the individual machine averages $\overline{Y}_i$,

$$\sum_{i=1}^{m} a_i \overline{Y}_i = a_1 \overline{Y}_1 + a_2 \overline{Y}_2 + \cdots + a_m \overline{Y}_m$$

with the weights summing to 1, that is,

$$\sum_{i=1}^{m} a_i = a_1 + a_2 + \cdots + a_m = 1$$

where $\overline{Y}_i = \sum_{j=1}^{n_i} Y_{ij}/n_i$ is the average of n_i measurements from the ith machine.

It is common to estimate μ by weighting each machine average based on the sample

size n_i

$$\hat{\mu}_0 = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n_i} Y_{ij}}{\sum_{i=1}^{m} n_i} = \frac{\sum_{i=1}^{m} n_i \overline{Y}_i}{\sum_{i=1}^{m} n_i} \tag{2}$$

so that the weights are proportional to the number of measurements used to calculate the machine average $\overline{Y}_i$. Alternatively, the unweighted estimate of μ

$$\hat{\mu}_1 = \frac{\sum_{i=1}^{m} \overline{Y}_i}{m} \tag{3}$$

assigns equal weights to each machine average $\overline{Y}_i$. When all n_i are equal, $\hat{\mu}_0$ and $\hat{\mu}_1$ are identical.

The third estimator of μ uses weights that are proportional to the inverse of the variance of $\overline{Y}_i$. The variance of $\overline{Y}_i$ is obtained by utilizing the assumption that machine biases are random. Based on this assumption, we know that individual machine averages $\overline{Y}_i$ are normally distributed with average μ and variance $\sigma_b^2 + \sigma_i^2/n_i$. In other words, the variance of $\overline{Y}_i$ is the sum of the variance due to machine differences and random error variance. This particular estimator of μ

$$\hat{\mu} = \sum_{i=1}^{m} \lambda_i \overline{Y}_i \tag{4}$$

where

$$\lambda_i = w_i / \sum_{j=1}^{m} w_j, \quad \textit{and} \quad w_i = 1/\mathrm{var}\left(\overline{Y}_i\right) = \left(\sigma_b^2 + \sigma_i^2/n_i\right)^{-1}$$

has some desirable statistical properties [7]. However, since σ_b^2 and σ_i^2 are usually unknown in practice, we have to estimate them based on our sample data before we can calculate the weights w_i. The sample variance of the ith machine is

$$S_i^2 = \frac{\sum_{j=1}^{n_i}\left(Y_{ij} - \overline{Y}_i\right)^2}{n_i - 1}.$$

While computing S_i^2 is straightforward, it is not quite as easy to estimate the between–machine variance σ_b^2. Paule and Mandel [4, 10, 11] provide an iterative algorithm for estimating σ_b^2 and we denote their estimate by $S_{b_0}^2$. It is also possible to derive estimates of σ_b^2 analytically based on statistical theory [12]. In particular,

$$S_{b_1}^2 = \frac{\sum_{i=1}^{m} n_i\left(\overline{Y}_i - \hat{\mu}_0\right)^2 - \sum_{i=1}^{m}\left(1 - n_i/\sum_{j=1}^{m} n_j\right) S_i^2}{\sum_{i=1}^{m} n_i - \sum_{i=1}^{m} n_i^2/\sum_{i=1}^{m} n_i} \tag{5}$$

and

$$S_{b_2}^2 = \frac{1}{m-1}\sum_{i=1}^{m}\left(\overline{Y}_i - \hat{\mu}_1\right)^2 - \frac{1}{m}\sum_{i=1}^{m}\frac{S_i^2}{n_i} \tag{6}$$

correspond to the previously discussed weighted and unweighted consensus value estimators $\hat{\mu}_0$ and $\hat{\mu}_1$, respectively. Since the parameter σ_b^2 is always positive, $S_{b_1}^2$ and $S_{b_2}^2$ are replaced by 0 whenever they turn out to be negative. (Negative variance estimates can occur for a number of reasons. See [13] for more information.) The estimator based on weights proportional to the inverse of the variance of $\overline{Y}_i$ is

$$\hat{\mu}_2 = \sum_{i=1}^{m} \hat{\lambda}_i \overline{Y}_i \tag{7}$$

where

$$\hat{\lambda}_i = \hat{w}_i / \sum_{j=1}^{m} \hat{w}_j, \quad \hat{w}_i = \left(S_b^2 + S_i^2/n_i \right)^{-1}$$

when sample–based estimates are used in place of σ_b^2 and σ_i^2 in eq. (4). The quantity S_b^2 used to determine the weights is replaced by either $S_{b_0}^2$, $S_{b_1}^2$, or $S_{b_2}^2$. In our application, the choice of S_b^2 has very little effect on the calculated consensus value $\hat{\mu}_2$ because individual machine sample sizes, variances, and averages are fairly similar. One subtle difference among the three random–bias estimators of μ is that the weights used in $\hat{\mu}_2$ are random variables, that is, they depend on actual data, while the weights in $\hat{\mu}_0$ and $\hat{\mu}_1$ are nonrandom since they depend only on individual machine sample sizes.

The problem of estimating μ using these three weighted averages has been studied by many authors. For example, when all σ_i^2 are equal (the variances associated with each machine are the same) Cochran [6] recommends using $\hat{\mu}_0$ if the differences among machines is negligible (σ_b^2 is small), $\hat{\mu}_1$ if σ_b^2 is large relative to σ_i^2, and $\hat{\mu}_2$ in intermediate cases.

In a simulation study, Weiler and Culpin [7] compare the variances of $\hat{\mu}_0$, $\hat{\mu}_1$, and $\hat{\mu}_2$ for two special cases: 1) the machine variances are equal, $\sigma_i^2 = \sigma^2$, and 2) machines are identical, $\sigma_b^2 = 0$. For the first case ($\sigma_i^2 = \sigma^2$), the Weiler and Culpin study indicates that $\hat{\mu}_2$ has a smaller variance than $\hat{\mu}_0$ and/or $\hat{\mu}_1$, except when all n_i are equal. For the second case ($\sigma_b^2 = 0$), Weiler and Culpin provide a rule for deciding among $\hat{\mu}_0$, $\hat{\mu}_1$, and $\hat{\mu}_2$ based on prior information about σ_i^2 when $m = 2$. Thus, the guidelines for choosing among $\hat{\mu}_0$, $\hat{\mu}_1$, and $\hat{\mu}_2$ to estimate the true overall average μ are well defined when all σ_i^2 are equal; there are, however, no specific recommendations for the general case when the machine variances are not equal.

The three estimators ($\hat{\mu}_0$, $\hat{\mu}_1$, and $\hat{\mu}_2$) of μ considered thus far can be used when machine biases are thought to be random. We now consider an estimator of μ that is based on the fixed–bias assumption. The minimax estimator of μ proposed by

Eberhardt, Reeve, and Spiegelman [9] is given by

$$\hat{\mu}_3 = \sum_{i=1}^{m} c_i \overline{Y}_i \tag{8}$$

where the c_i weights are chosen to minimize the maximum average squared error of $\hat{\mu}_3$

$$\sum_{i=1}^{m} c_i^2 S_i^2/n_i + \left[\sum_{i=1}^{m} c_i(U_i - L_i)/2\right]^2$$

subject to the constraints

$$c_i \geq 0 \quad \text{and} \quad \sum_{i=1}^{m} c_i = 1.$$

The quantities L_i and U_i are the lower and upper bounds of bias b_i. The bias bounds L_i and U_i are not estimated from the data, but are *known quantities.* If the bias bounds are identical for all machines, $\hat{\mu}_3$ is simply a weighted average of $\overline{Y}_i$ with weights

$$c_i = \frac{(S_i^2/n_i)^{-1}}{\sum_{j=1}^{m} \left(S_j^2/n_j\right)^{-1}}.$$

Thus $\hat{\mu}_3 = \hat{\mu}_2$ for situations in which machines are the same ($S_b^2 = 0$) and bias bounds are equal.

Now that we have defined our four consensus values, we should also present their associated standard errors. Formulas for computing the standard errors of the four weighted averages are shown below. Details involving the derivation of the standard error formulas can be found in [7, 9, 10, 12].

$$s.e.(\hat{\mu}_0) = \frac{\sqrt{\sum_{i=1}^{m} n_i^2/\hat{w}_i}}{\sum_{i=1}^{m} n_i} \tag{9}$$

$$s.e.(\hat{\mu}_1) = \frac{\sqrt{\sum_{i=1}^{m} 1/\hat{w}_i}}{m} \tag{10}$$

$$s.e.(\hat{\mu}_2) \approx \frac{1}{\sqrt{\sum_{i=1}^{m} \hat{w}_i}} \tag{11}$$

$$s.e.(\hat{\mu}_3) \approx \sqrt{\sum_{i=1}^{m} c_i^2 S_i^2/n_i} \tag{12}$$

Consensus Values for Charpy Data

Since there are known strategies (under the random–bias assumption) for choosing a specific weighted average when all σ_i^2 are equal, we should determine whether our data have this property before computing the consensus value. We can use Levene's test [14] of homogeneity of variances to determine whether the equality of the machine variances is an acceptable assumption. The hypothesis of interest, called the null hypotheses, is that the machine variances are equal. (In statistics, a null hypotheses can only be rejected; we can never prove that the null hypothesis is true.) To perform Levene's test, we must first compute the test statistic F, which is given by

$$F = \frac{\sum_{i=1}^{m} n_i \left(\overline{Z}_i - \overline{Z}\right)^2 /(m-1)}{\sum_{i=1}^{m}\sum_{j=1}^{n_i} \left(Z_{ij} - \overline{Z}_i\right)^2 / \left(\sum_{i=1}^{m} n_i - m\right)} \tag{13}$$

where

$$\begin{aligned} Z_{ij} &= \left|Y_{ij} - \overline{Y}_i\right| \\ \overline{Z}_i &= \sum_{j=1}^{n_i} Z_{ij}/n_i \\ \overline{Z} &= \sum_{i=1}^{m}\sum_{j=1}^{n_i} Z_{ij} / \sum_{i=1}^{m} n_i. \end{aligned}$$

Next, we need to define the values of test statistic F for which we will reject the null hypotheses. If the null hypothesis is true and all the machine variances σ_i^2 are equal, our test statistic F of eq. (13) is distributed as an F distribution with $d_1 = m - 1$ and $d_2 = \sum_{i=1}^{m}(n_i - 1)$ degrees of freedom. Thus, the decision rule for a size α test is to *reject* the null hypothesis and conclude that machine variances are *not* equal, if

$$F > F_{1-\alpha:d_1,d_2}$$

where $F_{1-\alpha:d_1,d_2}$ is the $100(1-\alpha)$th percentile of the F distribution with d_1 and d_2 degrees of freedom. In statistical terms, the quantity α usually a small number such as 0.1 or 0.15, represents the probability of *incorrectly* rejecting the null hypothesis and concluding that the machine variances are not equal. Also, if we do not reject the null hypothesis (concluding there is insufficient evidence to indicate that machine variances are not equal), the common variance σ^2 can be estimated by the "pooled" estimator

$$S^2 = \frac{\sum_{i=1}^{m} (n_i - 1)\, S_i^2}{\sum_{i=1}^{m} n_i - m}.$$

The pooled variance estimator will be used in place of the individual machine variances S_i^2 when it is assumed that machine variances are equal. (It is common practice to assume variances are equal if we do not reject the null hypothesis in Levene's test; however, this does not prove that the variances are equal.)

Example 1. This example illustrates the application of Levene's test and the computation of the consensus value and its standard error for a particular pilot lot in the Charpy machine certification program. The data are from High energy pilot lot HH30. Table 1 contains the summary statistics for each of the three reference Charpy machines.

TABLE 1–Summary statistics for lot HH30.

Machine	n_i	$\overline{Y}_i$ (J)	S_i^2 (J^2)
1	32	93.9963	4.5544
2	33	92.7503	6.2459
3	32	94.5620	4.0058

Levene's test is performed first to determine whether we can assume that machine variances are equal. The test statistic F for testing $\sigma_i^2 = \sigma^2$ is 0.9916, and the quantity $F_{0.9:2,94}$ from our F table is 2.3599. Thus, we do not reject the hypothesis that all σ_i^2 are equal because the test statistic is less than $F_{0.9:2,94}$. When the machine variances are equal and the n_i differ only slightly, Weiler and Culpin [7] show that the unweighted average $\hat{\mu}_1$ is the "best" random–bias estimator of μ. The consensus value is then given by

$$\hat{\mu}_1 = (93.9963 + 92.7503 + 94.5620)/3 = 93.7695 \text{ J}.$$

To obtain the standard error of the consensus value, we first obtain the pooled estimate of σ^2

$$S^2 = (31 \times 4.5544 + 32 \times 6.2459 + 31 \times 4.0058)/94 = 4.9493 \text{ J}^2$$

and use this value in eq. (5) to obtain $S_{b_1}^2 = 0.7132$ J^2 (the iterative procedure of Paule and Mandel yields $S_{b_0}^2 = 0.7073$ J^2). The standard error of the consensus value is

$$\begin{aligned} s.e.(\hat{\mu}_1) &= \frac{1}{3}\sqrt{(1/\hat{w}_1) + (1/\hat{w}_2) + (1/\hat{w}_3)} \\ &= \frac{1}{3}\sqrt{(0.7132 + 4.9493/32) + (0.7132 + 4.9493/33) + (0.7132 + 4.9493/32)} \\ &= 0.5374 \text{ J}. \end{aligned}$$

If the fixed–bias assumption is more appropriate and if we have no prior knowledge about the bias for each machine, we assign identical lower and upper bias bounds (the

actual numbers are not important) for each machine. The weights c_i are obtained as

$$(c_1, c_2, c_3) = (0.3299, 0.3402, 0.3299)$$

and $\hat{\mu}_3 = 93.7590$ J with the standard error 0.2259 J. If the following bias bounds

$$\begin{aligned}(L_1, L_2, L_3) &= (-1.2, -0.9, -0.9)\\ (U_1, U_2, U_3) &= (1.2, 0.9, 0.9)\end{aligned}$$

are specified, the weights are

$$(c_1, c_2, c_3) = (0, 0.5077, 0.4923)$$

and $\hat{\mu}_3 = 93.6422$ J with the standard error 0.2759 J. The actual bias bounds for individual machines used at NIST to determine the consensus value of a lot of specimens are *unknown*. The bounds used in this example are arbitrary and were specified for illustrative purposes only.

This example demonstrates an interesting feature of the fixed–bias problem formulation: it is possible to obtain zero weights. In other words, the zero weight of c_i means that the data from this machine are not used to determine the consensus value. From a statistical standpoint, it is never wise to throw away data unless there is a physical or engineering justification, so the possibility of obtaining zero weights is disturbing. However, the developers of the technique do not consider zero weights to be a problem, but rather an indicator that the particular machine associated with the zero weight is "relatively poor for determining the true value" [9]. Ultimately, the interpretation of zero weights is application dependent.

Example 2. The Low energy pilot data, summarized in Table 2 for each reference machine, demonstrates the performance of the four μ estimators when the data is unbalanced and machine variances are unequal.

TABLE 2–Summary statistics for lot LL34.

Machine	n_i	$\overline{Y}_i$ (J)	S_i^2 (J^2)
1	42	19.1751	0.8807
2	25	17.6799	0.3416
3	40	17.8087	0.3870

Table 2 shows that the data are not nearly as balanced as in lot HH30 since n_2 is much smaller than n_1 and n_3. Levene's test statistic for testing $\sigma_i^2 = \sigma^2$ is 5.9080, while

the value of $F_{0.9:2,104}$ is 2.3544. Therefore, we reject the null hypothesis and conclude that the machine variances are not equal since the test statistic is larger than $F_{0.9:2,104}$. Under these circumstances, there is no rule for choosing among the possible consensus values when machine biases are random. However, by examining the weights used to compute the various estimators, we can decide which random–bias estimator is best for this data. Table 3 contains the values of $\hat{\mu}_i$, $s.e.(\hat{\mu}_i)$, and the weights used to calculate each $\hat{\mu}_i$. For completeness, noninformative, or equal, bias bounds were specified to calculate $\hat{\mu}_3$.

TABLE 3–Weighted averages and weights.

i	$\hat{\mu}_i$ (J)	$s.e.(\hat{\mu}_i)$ (J)	weights		
0	18.3149	0.5070	0.3925	0.2336	0.3738
1	18.2212	0.4957	0.3333	0.3333	0.3333
2	18.2170	0.4773	0.3303	0.3339	0.3358
3	18.0573	0.0668	0.2127	0.3264	0.4609

The weights assigned to each $\overline{Y}_i$ for computing $\hat{\mu}_0$ are proportional to n_i, so the weights shown in Table 3 for $\hat{\mu}_0$ are quite similar for Machines 1 and 3 since their sample sizes are nearly equal, and the weight assigned to Machine 2 is small relative to the other weights because the group sample size is small. Although unequal sample sizes are accounted for in the computation of $\hat{\mu}_0$, information on within–machine variation is ignored. The unweighted average $\hat{\mu}_1$ also ignores information on within–machine variation S_i^2, since each $\overline{Y}_i$ is assigned equal weight, 0.3333. The Paule-Mandel estimator $\hat{\mu}_2$ assigns weights to machine averages based on values of S_b^2 and S_i^2/n_i obtained from the data, so both the between– and within–machine variation, as well as the group sample size, are accounted for in the computation of $\hat{\mu}_2$. The weights associated with $\hat{\mu}_2$ in Table 3 are similar for all machines because the small variance of Machine 2 compensates for the small sample size.

Since the hypothesis of equal within–machine variance was rejected by Levene's test, it seems appropriate to use methods, such as the random–bias estimator $\hat{\mu}_2$, which take machine variability into account to calculate the consensus values. The fixed–bias estimator $\hat{\mu}_3$ determines the weights based only on values of S_i^2/n_i. Thus the weight for Machine 3 is the largest because it has small variance while Machine 1, which has relatively high variability, is assigned the smallest weight.

We can compare the performance of the random–bias estimators $\hat{\mu}_0$, $\hat{\mu}_1$, and $\hat{\mu}_2$ by simulating data that has the same n_i and similar magnitudes of the within– and between–machine variances as our LL34 pilot lot data. The simulation study shows

that $\hat{\mu}_2$ has smaller variance than $\hat{\mu}_0$, and that the variance of $\hat{\mu}_1$ is nearly as small as the variance of $\hat{\mu}_2$. Since the variance of $\hat{\mu}_2$ was the smallest, the simulation study supports our conjecture that machine variability should be considered when machine variances are not equal. While the simulation results are useful in examining a particular data configuration, they cannot be interpreted in a general sense. The estimator $\hat{\mu}_3$ was not included in the simulation study because it was derived based on different assumptions than the other estimators. Therefore, the standard error of the fixed–bias estimator $\hat{\mu}_3$ is not comparable to the standard errors of the three random–bias estimators.

Reference Values for Charpy Data

Although we have demonstrated the computation of consensus values using data from the Charpy machine certification program, estimating μ for the Charpy program is actually a reference value problem since a fixed set of reference machines constitutes the population of interest and the goal of the program is to estimate the mean of the three reference machines to compare to candidate machine means. We will discuss two methods of computing the reference value based on weighted and unweighted averages.

The fixed–bias consensus value developed by Eberhardt et al. is a weighted average that can be used to compute a reference value

$$\hat{\mu}_4 = \sum_{i=1}^{m} c_i \overline{Y}_i,$$

by specifying identical bias bounds so that the weights are

$$c_i = \frac{\left(S_i^2/n_i\right)^{-1}}{\sum_{j=1}^{m}\left(S_j^2/n_j\right)^{-1}}.$$

If the individual machine variances are assumed to be equal, they can be replaced with the pooled variance S^2 so the weights would only be dependent on the group sample sizes n_i.

In the Charpy machine certification program, the reference value is currently computed as the unweighted average of the three reference machine averages

$$\hat{\mu}_5 = \frac{\sum_{i=1}^{m} \overline{Y}_i}{m} \tag{14}$$

If individual sample sizes are equal and machine variances are assumed to be equal, then the weighted and unweighted averages will result in the same reference value.

While the formulas for computing unweighted average in the reference value problem in eq. (14) and in the consensus value problem in eq. (3) are identical, the standard errors of the two estimates are different since the scope of the consensus value problem is different from the scope of the reference value problem. The standard errors of $\hat{\mu}_4$ and $\hat{\mu}_5$

$$s.e.(\hat{\mu}_4) \approx \frac{1}{\sqrt{\sum_{i=1}^{m}(S_i^2/n_i)^{-1}}} \tag{15}$$

$$s.e.(\hat{\mu}_5) = \frac{\sqrt{\sum_{i=1}^{m}(S_i^2/n_i)}}{m} \tag{16}$$

are the same as for $s.e.(\hat{\mu}_2)$ and $s.e.(\hat{\mu}_1)$, respectively, when the machines are identical ($S_b^2 = 0$). The individual machine variances in eqs. (15) and (16) can be replaced with the pooled variance S^2, if we can assume that machine variances are equal.

Example 1. Computing the reference value for pilot lot HH30 is straightforward. Based on the results of Levene's test performed in the previous section, we will assume that the individual machine variances are equal, so the pooled variance, $S^2 = 4.9493$ J, is used to determine the weights

$$(c_1, c_2, c_3) = (0.3299, 0.3402, 0.3299)$$

and the reference value is given by

$$\hat{\mu}_4 = (0.3299 \times 93.9963) + (0.3402 \times 92.7503) + (0.3299 \times 94.5620) = 93.8946 \text{ J},$$

with standard error

$$\begin{aligned} s.e.(\hat{\mu}_4) &= \frac{1}{\sqrt{(S_1^2/n_1)^{-1} + (S_2^2/n_2)^{-1} + (S_3^2/n_3)^{-1}}} \\ &= \frac{1}{\sqrt{(4.9493/32)^{-1} + (4.9493/33)^{-1} + (4.9493/32)^{-1}}} \\ &= 0.2258 \text{ J}. \end{aligned}$$

The reference value computed as the unweighted average of the three reference machine averages is $\hat{\mu}_5 = 93.7695$ J which corresponds to $\hat{\mu}_1$ computed in the previous section. The standard error of $\hat{\mu}_5$ is

$$\begin{aligned} s.e.(\hat{\mu}_5) &= \frac{\sqrt{(S_1^2/n_1) + (S_2^2/n_2) + (S_3^2/n_3)}}{m} \\ &= \frac{\sqrt{(4.9493/32) + (4.9493/33) + (4.9493/32)}}{m} \\ &= 0.2259 \text{ J}. \end{aligned}$$

Weighting the machine averages does not appear to affect the reference value significantly for this pilot lot probably because the pooled variance was used and the sample sizes were similar for the three machines. The standard errors of the two reference values are nearly identical as well.

Example 2. The Low energy pilot lot, LL22, was found to have significantly different variances among reference machines, so we will use individual machine variances rather than the pooled variance in the calculations. Since the individual machine variances and sample sizes are different, we expect the weighted average to be slightly more precise (smaller standard error) than the unweighted average.

The weights used to compute $\hat{\mu}_4$

$$(c_1, c_2, c_3) = (0.2127, 0.3264, 0.4609)$$

are identical to the weights used to determine the fixed–bias consensus value $\hat{\mu}_3$ in the previous section, so the weighted reference value, $\hat{\mu}_4 = 18.0573$ J and its standard error, $s.e.(\hat{\mu}_4) = 0.0668$ J also correspond to the previously computed values. The unweighted reference value $\hat{\mu}_5 = 18.2212$ J, has a larger standard error, $s.e.(\hat{\mu}_5) = 0.0702$ J, than the weighted reference value as expected.

In the Charpy program, most pilot lots have similar individual machine sample sizes and variances, so the unweighted and weighted reference values would produce virtually the same result.

Discussion

We have defined consensus values and reference values with respect to their model assumptions. Four methods for estimating the consensus value and two methods for computing the reference value were outlined, and their application was illustrated for Charpy machine certification program data.

A general recommendation regarding the use of any particular consensus value is not possible because one estimator is not uniformly better than the rest. Thus we can only examine the relative performance of the estimators and choose one that is adequate for the problem at hand. The unweighted consensus value $\hat{\mu}_1$ is probably satisfactory for many applications in which group sample sizes and variances are similar. However, it is also beneficial to compute the other types of estimators because the unweighted average may be less efficient in cases where the data is very unbalanced and/or the group variances are not equal. For example, if $\hat{\mu}_1$, $\hat{\mu}_2$ and their associated

weights are similar, we have additional assurance that the unweighted average $\hat{\mu}_1$ is performing adequately. In situations where some information is known regarding the physics of the testing methods such that bias bounds can be derived, the fixed–bias consensus value $\hat{\mu}_3$ may be optimal, especially when the number of machines is small.

Acknowledgments

The authors thank Dom Vecchia and Chris McCowan of the National Institute of Standards and Technology, and Hari Iyer of Colorado State University for their input during the preparation of this document. This work is a contribution of the National Institute of Standards and Technology and is not subject to copyright in the United States.

References

[1] ASTM E 1271-88, "Standard Practice for Qualifying Charpy Verification Specimens of Heat-Treated Steel, " pp. 1–7.

[2] ASTM E 23–88, "Standard Test Methods for Notched Bar Impact Testing of Metallic Materials," pp. 198–213.

[3] Splett, J. D. and Wang, J. C.-M., "Proposed Changes to Charpy V-Notch Machine Certification Requirements," *Pendulum Impact Machines: Procedures and Specimens for Verification, ASTM STP 1248*, T. A. Siewert and A. K. Schmieder, Eds., American Society for Testing and Materials, Philadelphia, 1995, pp. 182–194.

[4] Mandel, J. and Paule, R. C., "Interlaboratory Evaluation of a Material with Unequal Numbers of Replicates," *Analytical Chemistry*, **42**(11), 1970, pp. 1194–1197.

[5] Schiller, S. B. and Eberhardt, K. R. "Combining Data from Independent Chemical Analysis Methods," *Spectrochimica Acta*, **46B**, 1991, pp. 1607–1613.

[6] Cochran, W. G., "The Combination of Estimates From Different Experiments," *Biometrics*, **10**, 1954, pp. 101–129.

[7] Weiler, H. and Culpin, D., "On the Variance of Weighted Means," *Technometrics*, **12**(4), 1970, pp. 757–773.

[8] Birkes, D., Seely, J. and Azzam, A.-M., "An Efficient Estimator of the Mean in a Two–Stage Nested Model," *Technometrics*, **23**(2), 1981, pp. 143–148.

[9] Eberhardt, K. R., Reeve, C. P. and Spiegelman, C. H., "A Minimax Approach to Combining Means, With Practical Examples," *Chemometrics and Intelligent Laboratory Systems*, **5**, 1989, pp. 129–148.

[10] Paule, R. C. and Mandel, J. "Consensus Values and Weighting Factors," *Journal of Research of the National Bureau of Standards*, **87**(5), 1982, pp. 377–385.

[11] Paule, R. C. and Mandel, J., "Consensus Values, Regressions, and Weighting Factors," *Journal of Research of the National Institute of Standards and Technology*, **94**(3), 1989, pp. 197–203.

[12] Rao, P. S. R. S. and Sylvestre, E. A., "ANOVA and MINQUE Type of Estimators for the One–Way Random Effects Model," *Communication in Statistics – Theory and Methods*, **13**(14), 1984, pp. 1667–1673.

[13] Searle, S. R., Casella, G. and McCulloch, C. E., *Variance Components*, John Wiley & Sons, Inc., New York, 1992.

[14] Levene, H., "Robust Tests for Equality of Variances," *Contributions to Probability and Statistics: Essays in Honor of H. Hotelling*, I. Olkin, Ed., Stanford University Press, 1960.

Jolene D. Splett[1] and C. Ming Wang[1]

PROPOSED CHANGES TO CHARPY V–NOTCH MACHINE CERTIFICATION REQUIREMENTS

REFERENCE: Splett, J. D., and Wang, C. M., "Proposed Changes to Charpy V–Notch Machine Certification Requirements," Pendulum Impact Machines: Procedures and Specimens for Verification, ASTM STP 1248, Thomas A. Siewert and A. Karl Schmieder, Eds., p. 182, American Society for Testing and Materials, Philadelphia, 1995.

ABSTRACT: In 1989 the administration of the Charpy V–Notch Certification Program was assumed by the National Institute of Standards and Technology. The United States Army originated the program to insure the measurement integrity of Charpy V–notch machines across the country. The program has been operated for many years using candidate machine acceptance limits which can possibly be traced to a 1955 paper by Driscoll, however, the original statistical justification for using these acceptance criteria has been lost or never existed. A statistical analysis of recent certification program data indicates that the existing candidate machine acceptance limits should be modified. In this paper, we will discuss and justify potential changes to candidate machine acceptance limits.

KEYWORDS: notched–bar testing, reference specimens, Charpy V–notch machine certification program, pendulum impact machines, impact testing

---

The Charpy V–Notch Machine Certification Program has been operating for many years to ensure the measurement quality of Charpy machines across the country. Basically, the program works as follows. The National Institute of Standards and Technology (NIST) obtains a pilot lot of 100 specimens from a supplier and measures the

---

[1]Mathematical Statistician, Statistical Engineering Division, Information Technology Laboratory, National Institute of Standards and Technology, Boulder, CO 80303.

impact toughness of the specimens using three reference machines. Impact toughness is measured as absorbed energy in the test. If the 100 measurements meet certain criteria, then the remainder of the lot is machined and sent to NIST where an additional 30 specimens, called the confirm lot, are randomly selected and broken. If the absorbed energy of the confirm lot is in agreement with the absorbed energy of the pilot lot, the lot is certified as a reference material by NIST. Sets of five reference specimens are then sold to companies that want to certify their candidate Charpy machine. The specimens are broken using the candidate machine and the broken specimens, along with their observed absorbed energies, are sent to NIST for analysis. If the specimens and measurements are satisfactory, then the candidate Charpy machine can be certified.

The origins of the current Charpy V–Notch Machine Certification Program can be traced to a 1955 paper by Driscoll [1]. Evidently, the reliability of Charpy measurements was in question and people were reluctant to use Charpy machines for acceptance testing, so Driscoll conducted a large–scale study to prove that Charpy machines were reliable. Driscoll realized that the key to demonstrating reliability and reproducibility for this type of destructive testing was in the homogeneity of the material used in the study, so he was able to show that the Charpy machines were in fact reliable by carefully selecting homogeneous specimens. Driscoll's contribution to the certification program is observed in the general limits which encompass the majority of absorbed energy averages from samples of size five for the Low and High energy ranges. Specifically, most of the Low energy samples had an average absorbed energy within 1.4 Joules (J) of the lot average, while the average absorbed energy of five High energy specimens was usually within 5% of the lot average. These limits are identical to those used in the current certification program to pass or fail a company's Charpy machine.

Various aspects of the current program merit careful study if we are to understand and improve the performance of the certification process. For instance, the number of specimens in both the pilot lot and the candidate machine's verification set, and the definition of a single reference value from measurements on two or more "standard" machines, ultimately affect the error rates associated with classifying machines as conforming or nonconforming. In this article, however, we limit our attention to the properties of the current criteria by which machines are judged acceptable.

The main concern regarding the present certification criteria is the arbitrary nature of the acceptance limits. The justification for using these limits is unknown and does not appear to be linked to any statistical reasoning or evaluation of their performance. We intend to demonstrate the statistical properties of the existing limits as well as illustrate potential new candidate machine certification limits which are more objective and statistically defensible.

ASSUMPTIONS AND DEFINITIONS

Several assumptions regarding the certification program in general are necessary to simplify the discussion in this paper. While somewhat artificial, these assumptions nevertheless serve to illustrate and explain the potential shortcomings in current evaluation procedures.

For demonstration purposes, we will assume that the three reference machines used

to determine the reference value of the lot are basically identical. Under this assumption, the process of determining a reference value can be simplified, hypothetically, by assuming that all 100 pilot lot specimens are measured on one "standard" machine. Furthermore, we assume that the pilot lot is a representative sample of the machined lot of specimens. Thus the unknown average absorbed energy of all specimens in a lot can be estimated by breaking the random sample of 100 specimens known as the pilot lot. The computed average absorbed energy of the pilot lot serves as our estimate of the true absorbed energy of the lot.

Because of the destructive nature of the Charpy test, we are unable to separate the sources of variation inherent in testing. Specifically, the inherent variation among specimens in a lot cannot be distinguished from measurement error variation associated with operators, ambient conditions, and other noise sources. We will define the sum of these two sources of variation as the *system variation*, σ^2, which can be written as

$$\sigma^2 = \sigma_s^2 + \sigma_e^2 \tag{1}$$

where σ_s^2 is the true specimen variation and σ_e^2 is the true variation associated with measurement errors. Since specimen inhomogeneity is a critical component of the system variation, it is clear that one way to minimize σ^2 is to reduce the specimen variation.

Under our simplifying assumptions, the NIST system variance would be estimated by computing the sample variance of the pilot lot, S_p^2, while the candidate machine's system variation would be estimated by the sample variance of five reference specimens, S_c^2, measured on the machine under test. Since the pilot lot and the five reference specimens, or *verification set*, come from the same lot, the specimen variance, σ_s^2, is the same for every machine tested on a single lot. In addition, we will assume that both the absorbed energy and measurement error distributions can be well–approximated by Gaussian distributions, and that the variance of the candidate machine's measurement errors is the same as that of the NIST measurement system.

Finally, we make no provision for the occurrence of outliers in the calculations and results presented below. (In actual practice, a value is considered an outlier if there is physical evidence to indicate that a test result is "bad".) While outliers are an important issue, and their detection and treatment deserves more consideration, the only effect of generating outliers in our hypothetical situation would be to reduce the size of the verification set.

EXISTING LOW ENERGY CERTIFICATION LIMITS

Since the existing certification acceptance criterion are different for the Low and High energy ranges, we will discuss the two cases separately, starting with the Low energy specifications. A candidate Charpy machine is certified in the Low energy range if the measured values of five reference specimens meet two conditions: (i) the average of the five measurements is within $1.4J$ of the reference value, and (ii) the range of the five measurements is not greater than $2.8J$. (The range of a sample is the difference between the largest and smallest measurements.) If the five measurements

from the Charpy machine under test are denoted by C_1, C_2, C_3, C_4, C_5, then the Low energy certification criteria can be expressed as:

$$-1.4J \leq \overline{P} - \overline{C} \leq 1.4J \quad \text{and} \quad \max\{C_i\} - \min\{C_i\} \leq 2.8J \tag{2}$$

where $\overline{P}$ denotes the average of the 100 pilot lot specimens, and $\overline{C}$ is the average of the five reference specimens measured on the Charpy machine under test.

To understand the implications of the current Low energy certification criteria, we computed the probability of certifying a candidate machine based on the limits in (2) under the assumptions outlined in the previous section. The probability of certification depends on three parameters: the unknown average absorbed energy of the lot as defined by the NIST system, μ_p, the unknown average absorbed energy of the lot according to the candidate machine, μ_c, and the system standard deviation, σ, which is assumed to be the same for both NIST and the candidate machine. In other words, we calculated the probability that a machine will be certified (that is, the probability of observing $|\overline{P} - \overline{C}| \leq 1.4J$ and range$\{C_i\} \leq 2.8J$) for specified values of the parameters μ_p, μ_c, and σ. By specifying the parameters, we know a priori if the "simulated" candidate machine is behaving properly or not, but random variation makes it impossible to correctly classify a candidate machine as "good" or "bad" all the time. The probability of certification is given by

$$\Pr\left[-1.4J \leq \overline{P} - \overline{C} \leq 1.4J, \quad \text{range}\{C_i\} \leq 2.8J\right] =$$
$$5\left\{\Phi\left(\frac{1.4-\delta}{\sigma\sqrt{0.21}}\right) - \Phi\left(\frac{-1.4-\delta}{\sigma\sqrt{0.21}}\right)\right\} \cdot \int_{-\infty}^{\infty}\left\{\Phi\left(x + \frac{2.8}{\sigma}\right) - \Phi(x)\right\}^4 \phi(x)\,dx$$

where $\delta = \mu_p - \mu_c$, $\phi(x) = (2\pi)^{-1/2}e^{-x^2/2}$ is the standard Gaussian probability distribution, and $\Phi(x) = \int_{\infty}^{x}\phi(z)\,dz$. The probability of certification, calculated using numerical integration routines in IMSL [2], is shown in Fig. 1 as a function of $|\delta| = |\mu_p - \mu_c|$ and σ. (The probability is a symmetric function of δ.)

The horizontal axis in Fig. 1 represents the magnitude of the difference between the *true* average absorbed energy of specimens measured on a single candidate machine and specimens measured on the NIST reference machine. The vertical line at $1.4J$ in Fig. 1 is added as a point of reference; it denotes the point at which the difference between NIST and candidate machine *sample* average absorbed energies separates "good" and "bad" candidate machines according to the current certification criteria. Because of system variation, a "good" machine may fail, or a "bad" machine may pass purely by chance, and the probability of either type of error can be read from Fig. 1. For example, if $|\mu_p - \mu_c| = 0$ and the system standard deviation is high ($\sigma = 1.0J$), then the chance of correctly certifying a candidate machine is only about 72%, even though the candidate machine is in agreement with the NIST system.

If the current criterion $|\overline{P} - \overline{C}| \leq 1.4J$ is interpreted to mean that we would like to correctly certify candidate machines 100% of the time if $|\mu_p - \mu_c| \leq 1.4J$ while correctly failing candidate machines 100% of the time if $|\mu_p - \mu_c| > 1.4J$, then the ideal probability curve for certification limits would be a step function. Of course, perfect certification is possible only in the unrealistic circumstance that $\sigma = 0$. The

curves displayed in Fig. 1 show the chance of making an error, either by certifying a "bad" machine or not certifying a "good" machine, for three realistic values of the system standard deviation σ. The increasing frequency of misclassification as the system standard deviation increases from the smallest value, $\sigma = 0.6J$, to the largest standard deviation, $\sigma = 1.0J$, demonstrates that the performance of the certification criteria is very sensitive to random variation. We will return to this point shortly.

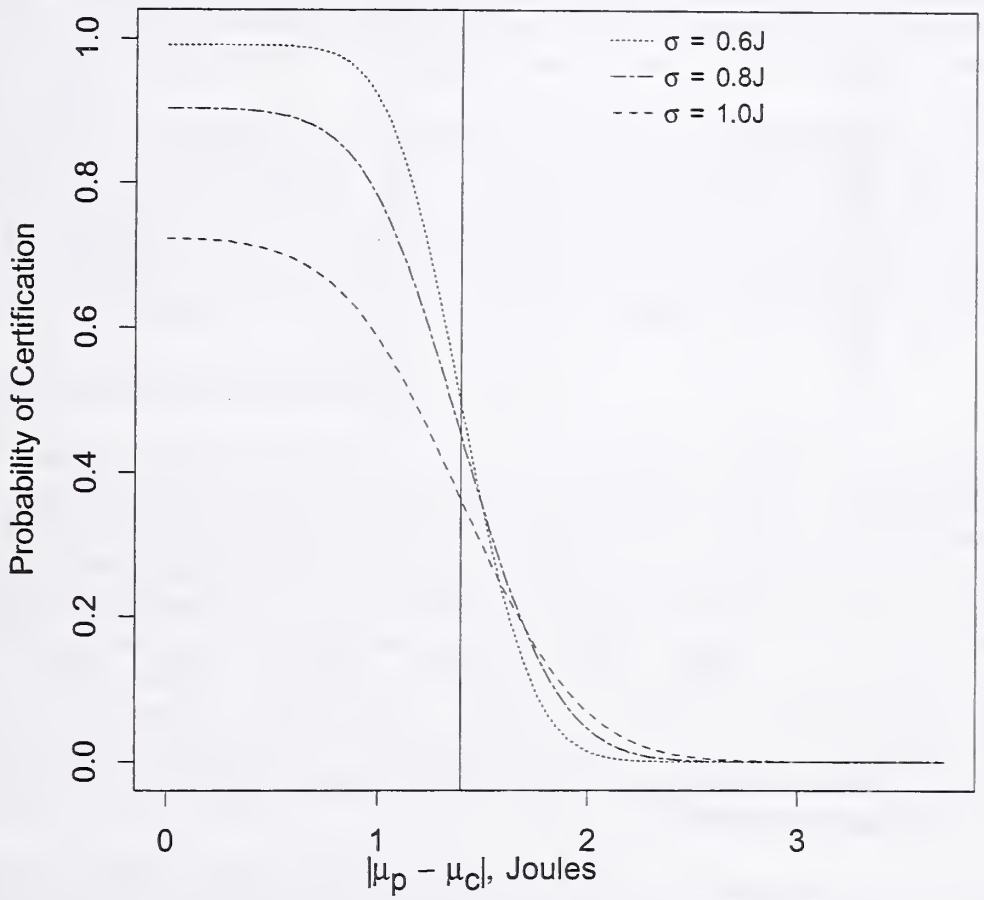


FIG. 1–Probability of certifying a candidate machine under the existing Low energy certification criteria.

EXISTING HIGH ENERGY CERTIFICATION LIMITS

To certify a candidate Charpy machine in the High energy range, the average of the five reference specimens measured on the candidate machine must be within 5% of the pilot lot average, or reference value; that is, we require $0.95\overline{P} \leq \overline{C} \leq 1.05\overline{P}$. In other words, there is only one certification limit which we may write as:

$$0.95 \leq \frac{\overline{C}}{\overline{P}} \leq 1.05 \tag{3}$$

where $\overline{P}$ and $\overline{C}$ are defined as in the Low energy case.

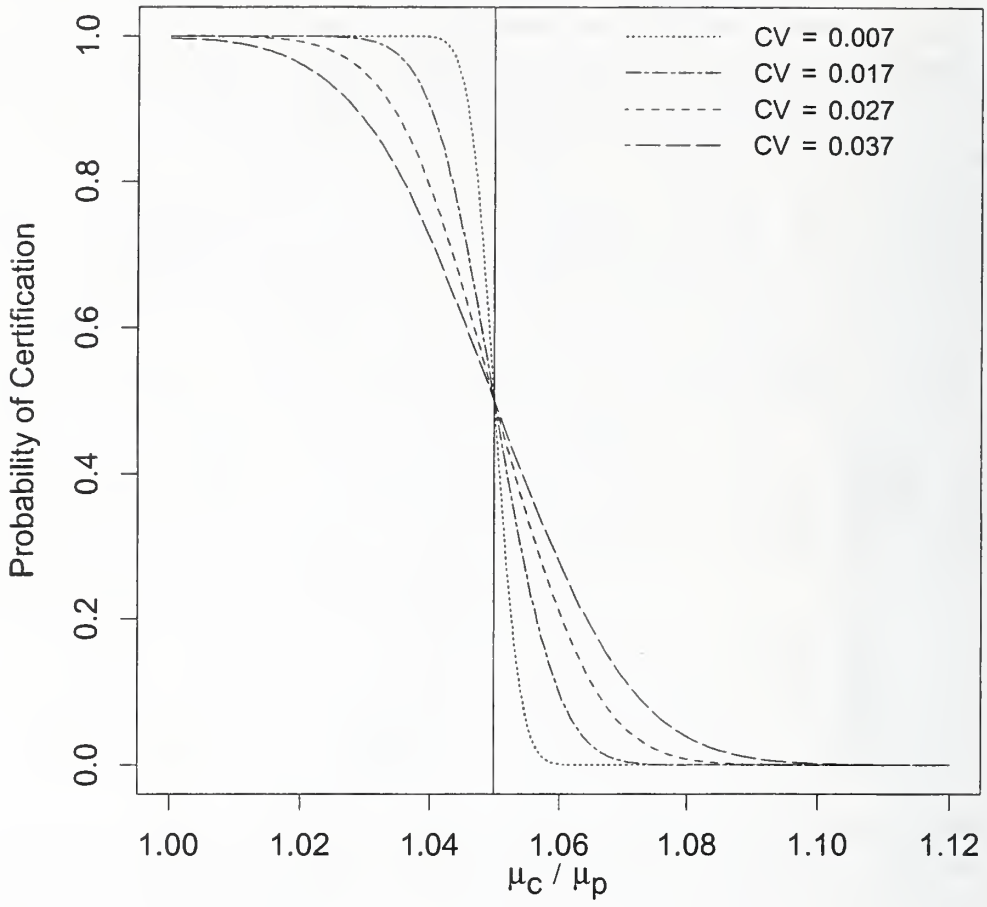


FIG. 2–Probability of certifying a candidate machine based on the existing High energy certification criterion.

To illustrate the properties of the current High energy certification criterion (3), we calculated the probability of certification for given values of the parameters μ_p, μ_c, and σ. The parameters are defined as in the Low energy case. The probability of certification for the High energy criterion is given by

$$\Pr\left[0.95\overline{P} \le \overline{C} \le 1.05\overline{P}\right] = \int_{-\infty}^{\infty} \left\{ \Phi\left(\frac{1.05\sqrt{5}x}{10} - \frac{\mu_c/\mu_p - 1.05}{\mathrm{CV}/\sqrt{5}} \right) - \Phi\left(\frac{0.95\sqrt{5}x}{10} - \frac{\mu_c/\mu_p - 0.95}{\mathrm{CV}/\sqrt{5}} \right) \right\} \phi(x)\, dx$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ were defined in the previous section. While there are three parameters, the certification probability only depends on the ratio, μ_c/μ_p, and the coefficient of variation, $CV = \sigma/\mu_p$, of the NIST reference machine. CV is a unitless measure of relative variation which may as well be called the noise–to–signal ratio. Fig. 2 displays

the probability of certification for the High energy range as a function of μ_c/μ_p for various values of CV that are acceptable according to the specimen manufacturer's contract. Since the probability of certification is symmetric about the value $\mu_c/\mu_p = 1$, it is sufficient to show the curves for ratios greater than or equal to 1.

The horizontal axis in Fig. 2 denotes the ratio of true averages, so machines having a ratio smaller than 1.05 could be called "good" while a "bad" machine would have a ratio larger than 1.05. From the figure, we see that the probability of certifying a candidate machine is roughly 0.5 when $\mu_c/\mu_p = 1.05$, regardless of the value of CV. In other words, if the average absorbed energy of the machine under test exceeds the NIST system average by 5%, the candidate machine essentially has a 50% chance of being certified.

As in the Low energy case, the probability of making an error deteriorates rapidly with increasing system variation. Fig. 2 shows how the probability of making an error, either by certifying a "bad" machine or not certifying a "good" machine, increases with CV. For example, if $\mu_c/\mu_p = 1.06$ the probability of incorrectly certifying a candidate machine is roughly 0.1 when $CV = 0.017$; however, if $CV = 0.037$ (i.e., a greater noise–to–signal ratio), the probability of incorrect certification rises to about 0.3.

ALTERNATE CERTIFICATION LIMITS

The results illustrated in Figs. 1 and 2 show that system variation, which comprises both inherent specimen–to–specimen differences as well as laboratory measurement errors, is the critical determinant of misclassification rates in the Charpy testing program. Yet, estimates of system variation (both for NIST and the candidate machine) are not assigned a prominent role in the existing certification procedure. Currently, the only part of the procedure which addresses system variation explicitly is the Low energy bound on the range of the candidate machine's verification set measurements. The range–test is designed to detect excessive candidate machine system variance, but does not actually incorporate any of the information from the pilot lot. Similarly, neither the Low or High energy criteria for comparing the average of the verification set to the reference value presently depend on the estimated NIST system variance, our best indicator of variability in the *current lot* of specimens.

An alternative to the current practice is to replace the existing Low and High energy certification procedures by a single protocol that incorporates system variation into the test procedure in two ways. First, the candidate machine's system variation would be compared to the NIST system variation by conducting a standard statistical test based on the sample variance of the verification set and the sample variance of the pilot lot. Because specimens come from the same lot, candidate machine variance which exceeds the lot variance can only be attributed to excessive measurement error variability in the candidate machine. The second way of accounting for system variation in the test criteria is to evaluate the difference between the reference value and the candidate machine's verification set average based on an acceptance limit which implicitly depends on the actual specimen–to–specimen variation of the current lot, so the computed limit would fluctuate with an estimated system variation based on the pilot lot and candidate machine data. A test for excessive candidate machine system

variability and alternate limits on the average of the verification set measurements are discussed below.

The standard test to compare NIST and candidate machine system variation is the ratio of the sample variance of the five reference specimens to the sample variance of the pilot lot. In this application we will only be concerned if the variance of the five reference specimens is large compared to the variance of the pilot lot. Regardless of the energy level, the sample variance of the five reference specimens will be considered too large if the ratio

$$F = \frac{S_c^2}{S_p^2} \tag{4}$$

is greater than the $100(1-\alpha)$th percentile of the F distribution, denoted by $F_{1-\alpha:4,99}$. Here S_c^2 and S_p^2 denote, respectively, the sample variance of the five reference specimens broken on the candidate machine and the sample variance of the pilot lot. The values 4 and 99 appearing in (4) are the respective degrees of freedom for S_c^2 and S_p^2, and α is the probability of *erroneously* concluding that the system variation of the candidate machine exceeds that of the NIST system when it does not. Usually α is small; for example, if $\alpha = 0.05$ we would claim that the system variation associated with a candidate machine is excessive if the F–ratio is greater than the the threshold value $F_{0.95:4,99} = 2.46$. See [3] for more information about statistical tests for comparing variances.

If the candidate machine passes the variability test, we would assume the system variation of the candidate machine does not exceed that of NIST, and the second test would be applied. The procedure we have developed is similar to the current Low energy criterion, but is based on an explicit statistical test of the hypothesis that $|\mu_p - \mu_c| \leq d$. By the new test, we will conclude that the candidate machine may be certified if $\overline{P} - \overline{C}$ is in some interval, say (L, U). The acceptable deviation d, which represents the amount by which the conceptual lot average of the candidate machine is permitted to differ from the NIST lot average, can be adjusted according to program requirements. The acceptance limits L and U, derived in the Appendix, are given by

$$U = -L = (d + 0.4583 \cdot S \cdot t_{1-\alpha:103})J \tag{5}$$

where $S^2 = (99S_p^2 + 4S_c^2)/103$, $t_{1-\alpha:103}$ is the $100(1-\alpha)$th percentile of Student's t distribution with 103 degrees of freedom, and α is the probability of failing a "good" candidate machine.

The particular value of the allowable deviation, d, between the true reference average and that of the machine under test at a given energy level is a choice that is based on engineering judgement. For illustration, we will choose $d = 1.4J$ to coincide with the current Low energy acceptance criterion. (No value of d is defined under the current program for High energy.) The particular value $d = 1.4J$ is not necessarily the best Low energy limit for actual use in the certification program. Driscoll [1] implies that system variation is already accounted for by the value $1.4J$.

Having specified d, the certification limits in (5) are quite easy to calculate; all that is required is a t–table. Under the current program, for example, the Low energy limit

is $d = 1.4J$. Supposing that a machine under test has passed the variability test in (4), and that the pooled standard deviation is $S = 0.5J$, then the final certification criterion when $\alpha = 0.05$ would require that $|\overline{P} - \overline{C}| \leq 1.7803J$. The acceptance limit was calculated by substituting d, S and the tabled value $t_{0.95:103} = 1.6598$ in (5).

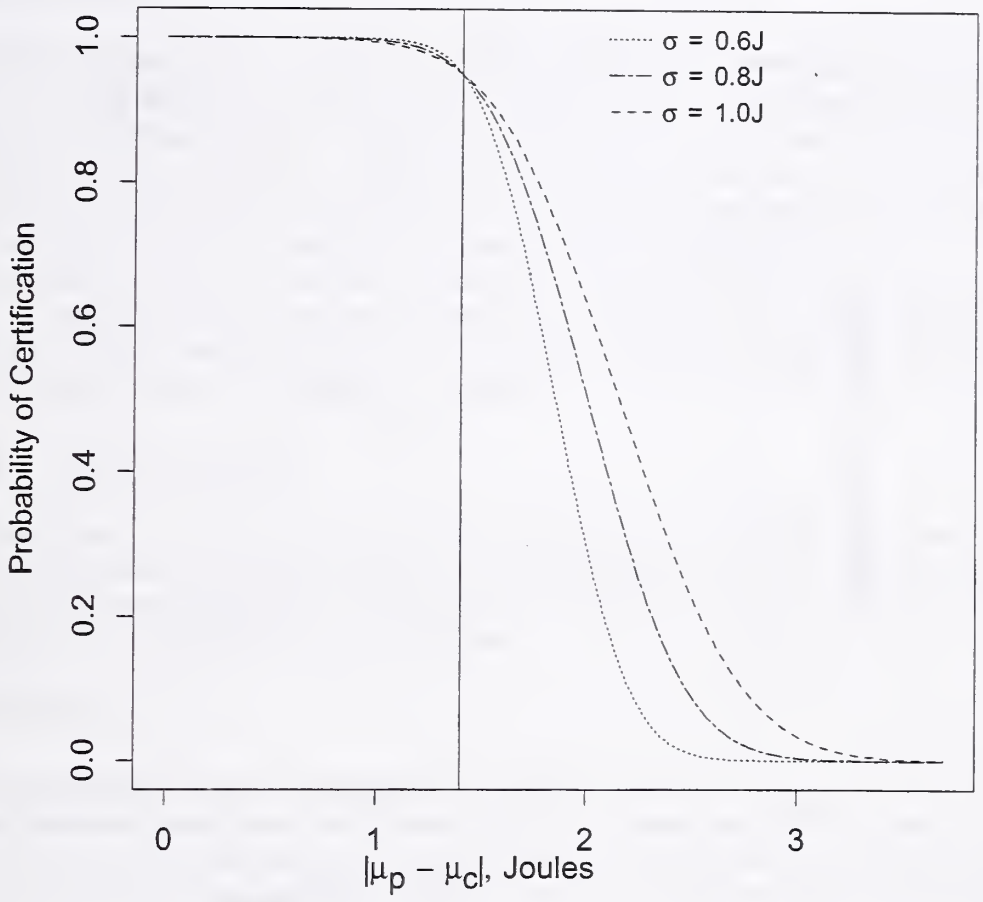


FIG. 3–Probability of certifying a candidate machine at the Low energy level versus the theoretical difference between the reference value and the true candidate machine average absorbed energy when $\alpha = 0.05$.

By contrast to the fixed limits on $\overline{P} - \overline{C}$ in (2), the new criterion based on (5) takes the variation of the specimens into consideration when computing the acceptance limits; the more variable the specimens, the wider the acceptance interval. The current Low energy acceptance interval $(-1.4J, 1.4J)$ is believed to be too stringent, resulting in unsatisfactorily low probabilities of certifying a "good" machine (see Fig. 1). The probability of certifying a candidate machine using the interval (L, U) based on $d = 1.4J$ is displayed in Fig. 3.

In Fig. 3, the probability of certification is quite close to 1.0 regardless of variation for values of $|\mu_p - \mu_c|$ less than $1.4J$; that is, a "good" machine will almost always

be certified. While the probability of certification is still dependent on σ, the main influence of increasing variability occurs when $|\mu_p - \mu_c| \geq 1.4J$, where the effect of increasing σ is to make it more likely that we will certify a "bad" machine. Again, a step function would be ideal, but progress toward the ideal could only be achieved by accepting lots with very small specimen variation.

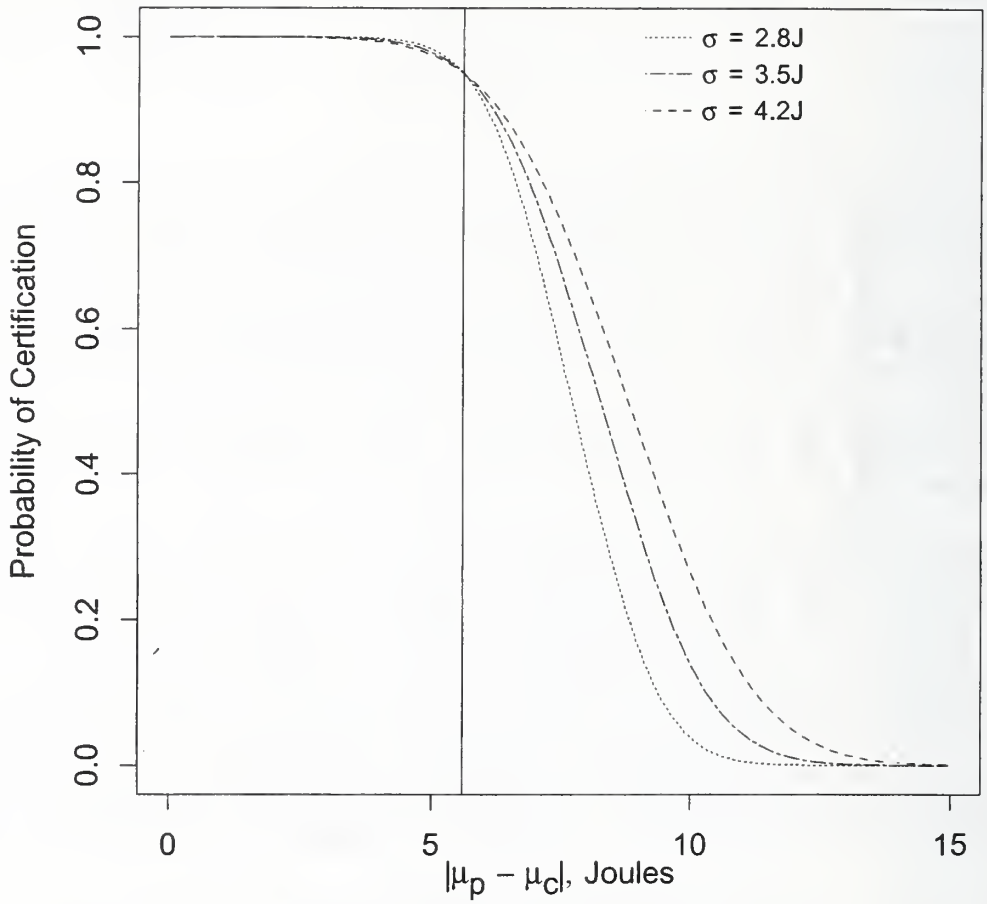


FIG. 4–Probability of certifying a candidate machine at the High energy level versus the theoretical difference between the reference value and the true candidate machine average absorbed energy when $\alpha = 0.05$.

The procedure described above can be applied to the High energy case as well. Assuming that the F–test has been passed, the acceptance limits for the High energy case would be calculated exactly as they were for the Low energy test, but substituting any appropriate value of d in (5). For illustrative purposes, we have used $d = 5.6J$ in (5) and computed the probability of certifying a candidate machine for realistic values of the variance for High–energy specimens. Fig. 4 displays the probability of certifying a candidate machine for the High energy case, assuming that the system variation for the candidate machine is the same as for NIST.

The probabilities shown in Fig. 4 are similar in structure to those observed for the Low energy case. The probability of certification is nearly 1.0 for "good" machines, $|\mu_p - \mu_c| \leq 5.6J$, while the probability of certifying "bad" machines, $|\mu_p - \mu_c| > 5.6J$, drops off gradually depending on σ.

CONCLUSIONS

We have shown that the current limits for certifying candidate Charpy machines need to be reconsidered, since they are not adjusted for lot–to–lot system variation. This paper introduced one alternate certification procedure that accounts for system variation and is more rigorous in its statistical validity. The procedure we have developed could be applied within the existing constraints of the program. There are many other potential procedures that could be considered for this problem.

Other aspects of the certification program that merit further study include: verification set sample size, pilot lot sample size, and a method for calculating appropriate estimates of the lot mean and NIST system standard deviation when there are differences among the three NIST reference machines.

ACKNOWLEDGEMENT

The authors thank Dom Vecchia, Chris McCowan, Dan Vigliotti and Tom Siewert of the National Institute of Standards and Technology for their input during the preparation of this document. This work is a contribution of the U. S. government, not subject to copyright in the United States.

APPENDIX

For the Lower energy, the hypothesis we want to test is $H_0 : |\mu_p - \mu_c| < 1.4J$. The decision rule is to reject H_0 if $\overline{P} - \overline{C}$ is not in $(L,\ U)$, where L and U are chosen so that the test has size α. The size of the test is defined as the probability of rejecting H_0 when H_0 is true, or the probability of not certifying a "good" machine. Notationally, we write

$$\Pr\left[\overline{P} - \overline{C} \notin (L,\ U) \mid -1.4J \leq \mu_p - \mu_c \leq 1.4J\right] = \alpha.$$

Common values of α used in practice are 0.05 and 0.01.

With a pre–determined α, acceptance limits L and U can then be determined from

$$\begin{aligned} 1 - \alpha &= \Pr\left[L \leq \overline{P} - \overline{C} \leq U \mid \mu_p - \mu_c = -1.4J\right] \\ &= \Pr\left[L \leq \overline{P} - \overline{C} \leq U \mid \mu_p - \mu_c = 1.4J\right] \end{aligned} \tag{6}$$

(e.g., see [4], p. 427). Equation (6) can be depicted in Fig. 5 if we assume that $\overline{P} - \overline{C}$ is normally distributed with average $\mu_p - \mu_c$ and variance $\sigma^2(1/100 + 1/5) = 0.21\sigma^2$.

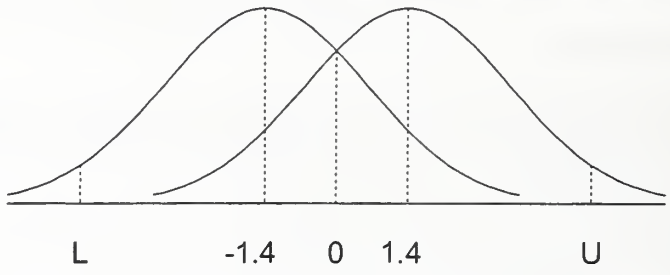


FIG. 5–Acceptance limits L and U with respect to two normal distributions.

From Fig. 5, it is easily seen that $U = -L$. Furthermore, by making use of the result that

$$\frac{\overline{P} - \overline{C} - (\mu_p - \mu_c)}{\sqrt{0.21}S}$$

is distributed as a Student's t with 103 degrees of freedom, the value of U satisfies

$$T_{103}\left(\frac{U + 1.4}{0.4583\,S}\right) - T_{103}\left(\frac{1.4 - U}{0.4583\,S}\right) = 1 - \alpha \tag{7}$$

where $S^2 = (99S_p^2 + 4S_c^2)/103$ is the pooled estimate for σ^2, and $T_n(\cdot)$ is the distribution function of Student's t with n degrees of freedom. If S is not too large, say $S \leq 4$, which should be the case in our problem, a simple solution for (7) is

$$U = 1.4 + 0.4583\,S\,t_{1-\alpha:103} \tag{8}$$

where $t_{1-\alpha:103}$ is the $1 - \alpha$ quantile of Student's t with 103 degrees of freedom. The solution can be verified as follows. With the solution in (8), the left–hand side of (7) reduces to

$$T_{103}\left(\frac{2.8}{0.4583\,S} + t_{1-\alpha:103}\right) - T_{103}\left(-t_{1-\alpha:103}\right) = T_{103}\left(\frac{2.8}{0.4583\,S} + t_{1-\alpha:103}\right) - \alpha.$$

With $S \leq 4$ and commonly used values of α,

$$T_{103}\left(\frac{2.8}{0.4583\,S} + t_{1-\alpha:103}\right)$$

is very close to 1, and (7) is satisfied.

REFERENCES

[1] Driscoll, D. E., "Reproducibility of Charpy Impact Test," *Symposium on Impact Testing, ASTM STP 176*, American Society for Testing and Materials, 1955, pp. 70–74.

[2] IMSL Inc., *The IMSL Math/Library*, Houston, Texas, 1987.

[3] Walpole, R. E. *Introduction to Statistics, Third Edition*, Macmillan Publishing Company, New York, 1982.

[4] Mood, A. M., Graybill, F. A, and Boes, D. C. *Introduction to the theory of Statistics, Third Edition*, McGraw–Hill Inc., New York, 1974.

[5] Fahey, N. H., "Effects of Variables in Charpy Impact Testing," *Materials Research and Standards*, Vol. 1, No. 11, November 1961, pp. 872–876.

[6] ASTM E 1236–91, "Standard Practice for Qualifying Charpy Impact Machines as Reference Machines," pp. 1–6.

[7] ASTM E 23–88, "Standard Test Methods for Notched Bar Impact Testing of Metallic Materials," pp. 198–213.

J. D. Splett[1] and C. M. Wang[1]

Uncertainty in Reference Values for the Charpy V-notch Verification Program

REFERENCE: Splett, J. D. and Wang, C. M., **"Uncertainty in Reference Values for the Charpy V-notch Verification Program,"** Journal of Testing and Evaluation, JTEVA, Vol. xx, No. x, December, 2002, pp. xx–xx.

ABSTRACT: We present a method for computing the combined standard uncertainty for reference values used in the Charpy machine verification program administered by the National Institute of Standards and Technology. The technique is compliant with the ISO GUM and models the between-machine bias using a Type B distribution. We demonstrate the method using actual data from the Charpy machine verification program.

KEYWORDS: Charpy V-notch, impact certification program, impact testing, ISO GUM, notched-bar testing, reference specimens, uncertainty

For the past 13 years, the National Institute of Standards and Technology (NIST) has administered a program to ensure the measurement integrity of Charpy V-notch machines across the nation [1]. A brief description of the program follows. NIST obtains a verification set of 75 impact specimens from a manufacturer and measures the impact toughness of each specimen on one of three "master" Charpy machines. Impact toughness is measured as energy in joules absorbed by the specimen during the test. If the verification set meets certain criteria, then the remaining specimens in the production lot will be machined. A sample of 15 specimens from the production lot is then tested on a single master machine to determine if the production lot is in agreement with the verification set. Once the production lot has been accepted, NIST assigns a reference value to the lot and sells sets of five specimens to companies who wish to certify their own Charpy machine. The program is administered within the guidelines of ASTM 1271-88 [2] and ASTM E23-88 [3].

Several other Charpy machine verification programs exist throughout the world, however they differ widely from the NIST program [4]. Since there are no international standard practices for verifying Charpy machines, it is important to develop some common ground for comparison. There is some interest in conducting a long-term interlaboratory comparison of Charpy machines using a master batch of specimens. To facilitate this comparison, a measure of the uncertainty in the computed reference value is needed. While other Charpy programs already utilize the uncertainty of the reference value, ASTM E23 does not provide guidelines for computing this quantity.

We propose a method for estimating the combined standard uncertainty in the computed reference value for the NIST Charpy machine verification program and demonstrate the method

[1] Statistical Engineering Division, National Institute of Standards and Technology, Boulder, CO 80305.

using actual data from the verification program. The method provides an uncertainty estimate that is compliant with NIST [5] and ISO GUM [6] guidelines.

Reference Value and Uncertainty

The reference value is based on verification set data and is defined as

$$\bar{X}_{\text{Ref}} = \frac{1}{3}\sum_{i=1}^{3}\bar{x}_i \ , \tag{1}$$

where $\bar{x}_i$ represents the average absorbed energy observed for the 25 specimens tested on the i^{th} master machine.

The combined standard uncertainty of the reference value can be determined by combining three components of standard uncertainty: within-machine standard uncertainty ($u(w)$), standard uncertainty due to machine bias ($u(b)$), and the standard uncertainty of specimen homogeneity ($u(h)$). The combined standard uncertainty (u_c) is

$$u_c = \sqrt{u^2(w)+u^2(b)+u^2(h)} \ . \tag{2}$$

The degrees of freedom associated with each of the three components of uncertainty ($\upsilon_w, \upsilon_b, \upsilon_h$) can be combined to obtain the effective degrees of freedom using the Welch-Satterthwaite formula [6],

$$\upsilon_{eff} = \frac{u_c^4}{\dfrac{u^4(w)}{\upsilon_w}+\dfrac{u^4(b)}{\upsilon_b}+\dfrac{u^4(h)}{\upsilon_h}} \ . \tag{3}$$

The effective degrees of freedom υ_{eff} associated with the total uncertainty u_c are used to determine the appropriate coverage factor for confidence intervals.

Within-Machine Standard Uncertainty

The within-machine standard uncertainty $u(w)$ is based on the "pooled" standard deviation S_P

$$S_P = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2+(n_3-1)s_3^2}{n_1+n_2+n_3-3}} \ , \tag{4}$$

where n_1, n_2, n_3 are the number of verification specimens tested on each of the three master machines and s_1, s_2, s_3 are the associated standard deviations. Typically 25 verification specimens are tested on each machine, so the within-machine standard uncertainty is

$$u(w) = \frac{S_P}{\sqrt{75}}. \tag{5}$$

The degrees of freedom υ_w associated with $u(w)$ are (25 + 25 + 25 - 3) = 72. Since Charpy testing is destructive, within-machine uncertainty and specimen inhomogeneity cannot be separated, so $u(w)$ contains both within-machine and specimen inhomogeneity.

Standard Uncertainty Due to Machine Bias

The standard uncertainty due to machine bias accounts for possible bias in the observed averages associated with each master machine. The value of $u(b)$ can be quantified using a technique that models the unknown biases with a Type B uncertainty distribution. (See Levenson et al. [7] for details regarding the technique.) Using observed data for 75 verification specimens (25 specimens tested on each of the three master machines), a rectangular distribution bounded by the extremes of the averages of the three master machines is used to model the machine biases so that $u(b)$ is

$$u(b) = \frac{|\bar{x}_{\min} - \bar{x}_{\max}|}{2\sqrt{3}}, \tag{6}$$

with degrees of freedom

$$\upsilon_b = \left(\frac{1}{2}\right)\frac{(\bar{x}_{\min} - \bar{x}_{\max})^2}{u^2(\bar{x}_{\min}) + u^2(\bar{x}_{\max})}. \tag{7}$$

The quantity $\bar{x}_{\min}$ corresponds to the smallest average among the three master machines and $u(\bar{x}_{\min})$ is the associated uncertainty of $\bar{x}_{\min}$. The largest average among the three master machines and the associated uncertainty are denoted by $\bar{x}_{\max}$ and $u(\bar{x}_{\max})$, respectively. (Note that $u(\bar{x}_{\min})$ and $u(\bar{x}_{\max})$ do not correspond to the smallest and largest uncertainties among the three machines. Rather, they are the uncertainties that correspond to the minimum and maximum averages.)

Standard Uncertainty Due to Specimen Inhomogeneity

The final component of standard uncertainty $u(h)$ can be thought of as a correction for specimen inhomogeneity and is based on test results for 25 verification specimens broken on a single master machine and the results for 15 production lot specimens tested on the same master machine. Let μ_0, σ_0 be the unknown true mean and standard deviation of absorbed energy (by a master machine) for the specimens in the verification lot, and μ_1, σ_1 the corresponding parameters in the production lot. Let $\bar{X}_0$, S_0, $\bar{X}_1$, and S_1 be the sample estimates for μ_0, σ_0, μ_1, and σ_1, respectively. We want to make inferences about μ_1 of the production lot based on the sample estimates $\bar{X}_0$, S_0 of the verification lot. That is, we want to find a standard

uncertainty S such that

$$\mathrm{P}[\bar{X}_0 - 2S \le \mu_1 \le \bar{X}_0 + 2S] \approx 0.95,$$

or

$$\mathrm{P}[\mu_1 - 2S \le \bar{X}_0 \le \mu_1 + 2S] \approx 0.95,$$

or

$$\mathrm{P}\left[\frac{\mu_1 - \mu_0}{\sigma_0/\sqrt{25}} - \frac{2S}{\sigma_0/\sqrt{25}} \le \frac{\bar{X}_0 - \mu_0}{\sigma_0/\sqrt{25}} \le \frac{\mu_1 - \mu_0}{\sigma_0/\sqrt{25}} + \frac{2S}{\sigma_0/\sqrt{25}}\right] \approx 0.95.$$

Assuming the distribution of $\dfrac{\bar{X}_0 - \mu_0}{\sigma_0/\sqrt{25}}$ is approximately standard normal, then

$$\mathrm{P}\left[-2 \le \frac{\bar{X}_0 - \mu_0}{\sigma_0/\sqrt{25}} \le 2\right] \approx 0.95.$$

If we know the true values of μ_0, μ_1, and σ_0, one can show that an appropriate choice for S is

$$S = \frac{\sigma_0}{5}\left(1 + \frac{|\mu_1 - \mu_0|}{2\sigma_0/5}\right).$$

Thus, the uncertainty $\sigma_0/5$ is inflated by a factor of $1 + \dfrac{|\mu_1 - \mu_0|}{2\sigma_0/5}$.

Substituting the sample estimates $\bar{X}_0$, $\bar{X}_1$, and $u(\bar{X}_0)$ from the verification lot data for the true values, the inflation factor becomes $1 + \dfrac{|\bar{X}_1 - \bar{X}_0|}{2u(\bar{X}_0)}$. Once the inflation factor is estimated, we can use the uncertainty information from all the verification lot data to obtain $u(h)$. That is,

$$\sqrt{u^2(w) + u^2(b)}\left(1 + \frac{|\bar{X}_1 - \bar{X}_0|}{2u(\bar{X}_0)}\right) = \sqrt{u^2(w) + u^2(b) + u^2(h)}.$$

Thus, if the production lot is accepted, $u(h)$ is calculated using

$$u(h) = \sqrt{(u^2(w) + u^2(b))\left[\left(1 + \frac{|\bar{X}_1 - \bar{X}_0|}{2u(\bar{X}_0)}\right)^2 - 1\right]}, \tag{8}$$

where $\bar{X}_1$ is the average absorbed energy of the 15 production lot specimens tested using a master machine, and $\bar{X}_0$ is the average absorbed energy of the 25 verification specimens tested on the same master machine. A conservative estimate of the degrees of freedom associated with S_h is $\upsilon_h = 15 - 1 = 14$. If the distributions of results for the 25 verification specimens and the 15

production lot specimens are in good agreement, then the standard uncertainty due to specimen inhomogeneity will be small.

Example

The quantitative measurement results for an actual verification set are shown in Table 1. Fig. 1 displays box plots of the verification set measurements for each master machine. Each box plot shows the minimum, 25th percentile, median, 75th percentile, and maximum impact energy observed for each machine.

Table 1 – Verification set measurement results.

Machine #1	Machine #2	Machine #3
$n_1 = 25$	$n_2 = 25$	$n_3 = 25$
$\bar{x}_1 = 219.782$ J	$\bar{x}_2 = 226.761$ J	$\bar{x}_3 = 226.408$ J
$s_1 = 7.488$ J	$s_2 = 6.077$ J	$s_3 = 6.669$ J
$u(\bar{x}_1) = 1.498$ J	$u(\bar{x}_2) = 1.215$ J	$u(\bar{x}_3) = 1.334$ J

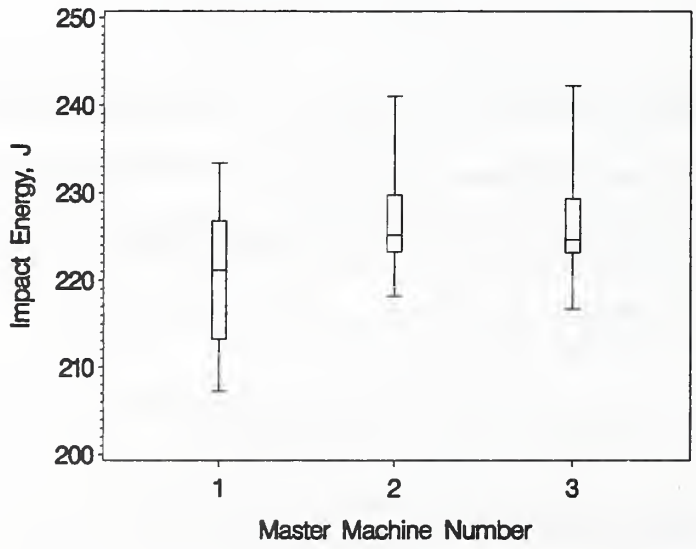


FIG. 1 – *Distribution of verification set measurements for each master machine.*

The reference value (see equation (1)) associated with the verification set shown in Table 1 is 224.317 J. To compute the combined standard uncertainty associated with the reference value,

we must first compute S_P using equation (4) and $u(w)$ from equation (5) (with $\upsilon_w = 72$ degrees of freedom).

$$S_P = \sqrt{\frac{s_1^2 + s_2^2 + s_3^2}{3}} = \sqrt{\frac{(7.488)^2 + (6.077)^2 + (6.669)^2}{3}} = 6.769 \text{ J}$$

$$u(w) = \frac{S_P}{\sqrt{75}} = \frac{6.769}{\sqrt{75}} = 0.782 \text{ J}$$

Next, equations (6) and (7) are used to compute $u(b)$ and υ_B, respectively, where $\bar{x}_{\min} = 219.782$ J, $\bar{x}_{\max} = 226.761$ J, $u(\bar{x}_{\min}) = 1.498$ J, and $u(\bar{x}_{\max}) = 1.215$ J.

$$u(b) = \frac{|\bar{x}_{\min} - \bar{x}_{\max}|}{2\sqrt{3}} = \frac{|219.782 - 226.761|}{2\sqrt{3}} = 2.015 \text{ J}$$

$$\upsilon_b = \left(\frac{1}{2}\right)\frac{(\bar{x}_{\min} - \bar{x}_{\max})^2}{u^2(\bar{x}_{\min}) + u^2(\bar{x}_{\max})} = \left(\frac{1}{2}\right)\frac{(219.782 - 226.761)^2}{(1.498)^2 + (1.215)^2} = 6.546,$$

and υ_B is rounded down to 6.

If 15 production lot specimens are tested on Machine #3 with $\bar{x}_{\text{Production}} = 223.738$ J, then from equation (8), $u(h)$ (with $\upsilon_h = 14$) is

$$u(h) = \sqrt{(u^2(w) + u^2(b))\left[\left(1 + \frac{|\bar{x}_{\text{Production}} - \bar{x}_3|}{2u(\bar{x}_3)}\right)^2 - 1\right]}$$

$$= \sqrt{((0.782)^2 + (2.015)^2)\left[\left(1 + \frac{|223.738 - 226.408|}{2(1.334)}\right)^2 - 1\right]}$$

$$= 3.745 \text{ J}.$$

The combined standard uncertainty (see equation (2)) associated with the reference value is

$$u_c = \sqrt{u^2(w) + u^2(b) + u^2(h)} = \sqrt{(0.782)^2 + (2.015)^2 + (3.745)^2} = 4.324 \text{ J}.$$

The degrees of freedom, calculated using equation (3), are

$$\upsilon_{eff} = \frac{u_c^4}{\frac{u^4(w)}{\upsilon_w} + \frac{u^4(b)}{\upsilon_b} + \frac{u^4(h)}{\upsilon_h}} = \frac{(4.324)^4}{\frac{(0.782)^4}{72} + \frac{(2.015)^4}{6} + \frac{(3.745)^4}{14}} = 20.805,$$

which rounds down to 20.

Thus, the expanded uncertainty, corresponding to a 95 % confidence interval on the true

reference value, is $t_{0.025,20}(4.324) = 2.086(4.324) = 9.020\,\text{J}$

Conclusions

We have presented a method for computing the combined standard uncertainty of a reference value for the Charpy machine verification program and have demonstrated the application of the method for actual data. The method is compliant with ISO GUM and NIST uncertainty guidelines.

By developing a procedure for computing the standard uncertainty of a reference value, we hope to provide a means for improving the limits used to certify customer Charpy machines. The certification limits currently in use are somewhat arbitrary and do not account for uncertainty in the reference value. The development of an uncertainty for a reference value has no practical effect on the verification program at this time.

Acknowledgements

The authors thank Chris McCowan of the National Institute of Standards and Technology for sharing his expertise regarding the Charpy program, and Hari Iyer of Colorado State University for his technical suggestions. This work is a contribution of the National Institute of Standards and Technology and is not subject to copyright in the United States.

References

[1] McCowan, C. N., Siewert, T. A., Vigliotti, D. P., The NIST Charpy V-notch Verification Program: Overview and Operating Procedures, NIST IR 6618, NIST, Gaithersburg, MD, in press.

[2] ASTM E 1271-88, "Standard Practice for Qualifying Charpy Verification Specimens of Heat-Treated Steel," American Society for Testing and Materials, Philadelphia.

[3] ASTM E 23-88, "Standard Test Methods for Notched Bar Impact Testing of Metallic Materials," American Society for Testing and Materials, Philadelphia.

[4] EN 10045-2: 1993, "Charpy Impact Test on Metallic Materials – Part 2: Method for the Verification of Impact Testing Machines," European Committee for Standardization, Brussels.

[5] Taylor, B. N., Kuyatt, C. E., Guidelines for Evaluating and Expressing Uncertainty in NIST Measurement Results, NIST TN 1297, NIST, Gaithersburg, MD, 1994.

[6] International Organization for Standardization (ISO), Guide to the Expression of Uncertainty in Measurement, International Organization for Standardization (ISO), Geneva, Switzerland, 1993 (corrected and reprinted 1995).

[7] Levenson, M. S., Banks, D. L., Eberhardt, K. R., Gill, L. M., Guthrie, W. F., Liu, H. K.,

Vangel, M. G., Yen, J. H., and Zhang, N. F., "An Approach to Combining Results From Multiple Methods Motivated by the ISO GUM," Journal of Research of the National Institute of Standards and Technology, Vol. 105, No. 4, July-August, 2000, pp. 571-579.

NIST Technical Publications

Periodical

Journal of Research of the National Institute of Standards and Technology—Reports NIST research and development in metrology and related fields of physical science, engineering, applied mathematics, statistics, biotechnology, and information technology. Papers cover a broad range of subjects, with major emphasis on measurement methodology and the basic technology underlying standardization. Also included from time to time are survey articles on topics closely related to the Institute's technical and scientific programs. Issued six times a year.

Nonperiodicals

Monographs—Major contributions to the technical literature on various subjects related to the Institute's scientific and technical activities.

Handbooks—Recommended codes of engineering and industrial practice (including safety codes) developed in cooperation with interested industries, professional organizations, and regulatory bodies.

Special Publications—Include proceedings of conferences sponsored by NIST, NIST annual reports, and other special publications appropriate to this grouping such as wall charts, pocket cards, and bibliographies.

National Standard Reference Data Series—Provides quantitative data on the physical and chemical properties of materials, compiled from the world's literature and critically evaluated. Developed under a worldwide program coordinated by NIST under the authority of the National Standard Data Act (Public Law 90-396). NOTE: The Journal of Physical and Chemical Reference Data (JPCRD) is published bimonthly for NIST by the American Institute of Physics (AIP). Subscription orders and renewals are available from AIP, P.O. Box 503284, St. Louis, MO 63150-3284.

Building Science Series—Disseminates technical information developed at the Institute on building materials, components, systems, and whole structures. The series presents research results, test methods, and performance criteria related to the structural and environmental functions and the durability and safety characteristics of building elements and systems.

Technical Notes—Studies or reports which are complete in themselves but restrictive in their treatment of a subject. Analogous to monographs but not so comprehensive in scope or definitive in treatment of the subject area. Often serve as a vehicle for final reports of work performed at NIST under the sponsorship of other government agencies.

Voluntary Product Standards—Developed under procedures published by the Department of Commerce in Part 10, Title 15, of the Code of Federal Regulations. The standards establish nationally recognized requirements for products, and provide all concerned interests with a basis for common understanding of the characteristics of the products. NIST administers this program in support of the efforts of private-sector standardizing organizations.

Order the **above** *NIST publications from: Superintendent of Documents, Government Printing Office, Washington, DC 20402 or* http://bookstore.gpo.gov/.

Order the **following** *NIST publications—FIPS and NISTIRs—from the National Technical Information Service, Springfield, VA 22161 or* http://www.ntis.gov/products.

Federal Information Processing Standards Publications (FIPS PUB)—Publications in this series collectively constitute the Federal Information Processing Standards Register. The Register serves as the official source of information in the Federal Government regarding standards issued by NIST pursuant to the Federal Property and Administrative Services Act of 1949 as amended, Public Law 89-306 (79 Stat. 1127), and as implemented by Executive Order 11717 (38 FR 12315, dated May 11, 1973) and Part 6 of Title 15 CFR (Code of Federal Regulations).

NIST Interagency or Internal Reports (NISTIR)—A special series of interim or final reports on work performed by NIST for outside sponsors (both government and nongovernment). In general, initial distribution is handled by the sponsor; public distribution is by the National Technical Information Service, Springfield, VA 22161 or http://bookstore.gpo.gov/index.html, in hard copy, electronic media, or microfiche form. NISTIR's may also report results of NIST projects of transitory of limited interest, including those that will be published subsequently in more comprehensive form.

U.S. Department of Commerce
National Institute of Standards and Technology
325 Broadway
Boulder, Colorado 80305-3337

Official Business
Penalty for Private Use, $300